

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

```
Run on:      October 7, 2001, 02:32:15 ; Search time 997.91 Seconds
              (without alignments)
              5712.009 Million cell updates/sec
```

```

Title:      US-09-735-712-1
Perfect score: 603
Sequence:   1 atggttcgaagcaccgcaca.....atgtgacaatgttttga 603

```

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1022815 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

```

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

```

Post-processing:	Minimum Match	0%
	Maximum Match	100%
	Listing first	45 summaries

Database :

EST: *

1:	qb.est1:	*
2:	qb.est2:	*
3:	qb.est3:	*
4:	qb.est4:	*
5:	qb.est5:	*
6:	qb.est6:	*
7:	qb.est7:	*
8:	qb.est8:	*
9:	qb.est9:	*
10:	qb.est10:	*
11:	qb.est11:	*
12:	qb.est12:	*
13:	qb.est13:	*
14:	qb.est14:	*
15:	qb.est15:	*
16:	qb.est16:	*
17:	qb.est17:	*
18:	qb.est18:	*
19:	qb.est19:	*
20:	qb.est20:	*
21:	qb.est21:	*
22:	qb.est22:	*
23:	qb.est23:	*
24:	qb.est24:	*
25:	qb.est33:	*
26:	qb.est34:	*
27:	qb.est35:	*
28:	qb.est36:	*
29:	qb.est37:	*
30:	qb.est38:	*
31:	qb.est39:	*
32:	qb.est40:	*
33:	em.est1:	*
34:	em.est1un1:	*
35:	em.est1un2:	*
36:	em.est1un3:	*
37:	em.est1un4:	*
38:	em.est1un5:	*
39:	em.est1un6:	*
40:	em.est1un7:	*
41:	em.est1un8:	*
42:	em.est1un9:	*

```
117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*
```

```
190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	466.2	77.3	508	16 A1149899	A1149899 qf43h06.x
C 2	450.6	74.7	516	7 AA436088	AA436088 zud3a08.r
C 3	347.7	57.5	387	7 AA416972	AA416972 zt94h05.s
C 4	340.2	56.4	382	7 AA435988	AA435988 zud3a08.s
C 5	291.8	48.4	394	11 AA758635	AA758635 ab67b04.s
C 6	283.8	47.1	484	163 BE107659	BE107659 U1-R-BT1-
C 7	267.4	44.3	410	7 AA411806	AA411806 zt67a03.s
C 8	245.4	40.7	389	11 AA781801	AA781801 a158q01.s
C 9	242.8	40.3	415	7 AA470059	AA470059 zt94h05.r
C 10	207	34.3	482	14 A1002083	A1002083 c038b02.s
C 11	141	23.4	332	147 BE319786	BE319786 uy63h11.x
C 12	131.8	21.9	601	138 BE638317	BE638317 EST00003
C 13	126.2	22.4	138	BE638325	BE638325 EST00022
C 14	115.6	19.2	538	224 AQ108532	AQ108532 CIT-HSP-2
C 15	86.4	18.7	281	11 AA707529	AA707529 ah41a12.s
C 16	86.4	14.3	1071	106 AL544561	AL544561 AL544561
C 17	83	13.9	793	154 BG484817	BG484817 602505511
C 18	84	13.8	218	138 BE638321	BE638321 EST00013
C 19	82.2	13.6	704	153 BE571626	BE571626 602592934
C 20	80.4	13.3	508	4 AA234138	AA234138 zt51b06.r
C 21	80.4	13.3	562	7 AA418443	AA418443 zt92e05.r
C 22	70.6	11.7	421	257 B86842	B86842 RPII11-26F1
C 23	61.2	10.1	887	106 AL531049	AL531049 AL531049
C 24	59.4	9.9	715	155 BG538851	BG538851 602568296
C 25	55.6	9.2	579	136 BE513276	BE513276 601315340
C 26	55.6	9.2	947	192 AK008652	AK008652 Mus muscu
C 27	55.4	9.2	470	102 A1807884	A1807884 wf43q11.x
C 28	55	9.1	528	138 BE675149	BE675149 7f03d02.x
C 29	55	9.1	904	172 BG024663	BG024663 602275469
C 30	54.6	9.1	687	143 BF056859	BF056859 7k10g10.x
C 31	54.4	9.0	1504	192 AK003110	AK003110 Mus muscu
C 32	53	8.8	1184	192 AK017928	AK017928 Mus muscu
C 33	52.6	8.7	515	153 BG434337	BG434337 602506331
C 34	52.6	8.7	594	32 AV717594	AV717594 AV717594
C 35	52.6	8.7	597	32 AV716310	AV716310 AV716310
C 36	52.6	8.7	619	32 AV661743	AV661743 AV661743
C 37	52.6	8.7	880	155 BG546745	BG546745 602574115
C 38	52.6	8.7	735	32 AV715678	AV715678 AV715678
C 39	52.4	8.7	504	150 BF522968	BF522968 U1-R-C2P-
C 40	52.4	8.7	755	104 A1950360	A1950360 wp10h06.x
C 41	52.2	8.7	736	137 BE563951	BE563951 601348160
C 42	52.2	8.7	860	106 AL551751	AL551751 AL551751
C 43	52	8.6	534	151 BF591089	BF591089 7h52g02.x
C 44	51.4	8.5	559	19 A1391038	A1391038 mc10h04.Y
C 45	50.8	8.4	482	151 BF600732	BF600732 265501.MA

ALIGNMENTS

RESULT 1	508 bp	EST	10-NOV-1998
A1149899/c	508 bp	EST	10-NOV-1998
DEFINITION	qf43h06.x1 Soares, testis_NHT Homo sapiens cDNA clone IMAGE:1752827		
LOCUS	3, similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ;		
ACCESSION	A1149899		
VERSION	A1149899		
KEYWORDS	mRNA sequence.		
SOURCE	A1149899.1 GI:3678368		
ORGANISM	human.		
REFERENCE	Mammalia: Eutheria: Primates: Catarrhini: Hominiidae: Homo.		
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/cgap.		
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
	Tumor Gene Index		

JOURNAL COMMENT

Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www.bio.lnl.gov/bdip/image/image.html
 Insert Length: 742 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 455.
 Location/Qualifiers
 1. 508
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1752827"
 /clone_1b="Soares, testis_NHT"
 /sex="male"
 /lab_host="DH10b"
 /note="Vector: pT7AD-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dt) primer [5', TCTTACCAATCGAAGTGGAGCGGCCCAATTTTCTTTTCTT 3']
 Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Col5, and was constructed by Bento Soares and M. Fatima Bonaldo."

FEATURES

source

BASE COUNT

186 a 91 c 89 g 142 t
 ORIGIN

Query Match 77.3% Score 466.2; DB 16; Length 508;
 Best Local Similarity 99.4% Pred. No. 5e-109;
 Matches 468; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	133	agaaaaatgaataatccttaggagacatcaagctctgtttggaattatgacctttctt	192
DB	508	AGAAAAATGAATAATCTTAGGACATCAAGTCTGTTGGAATATGACCTTTCTTTT	449
QY	193	ggaattatctctcttccatctgtttaaaccatcaaggttcccttataattctt	252
DB	448	GGAATTATCTCTCTTCCATCTTTGTTAAACATATCAAGTTCCCTTATTTCTT	389
QY	253	tcaagatataccatctcgggctcgtttgttcaatattcgtgagccttcaattgca	312
DB	388	TCAAGATATCCATCTCGGGCTCTGTTTGTTCATTAATCTGGAGCCTTCAATTGCA	329
QY	313	gtgaaaaagaataaacacagaacctcgataatattgagccgaataagaattctt	372
DB	328	GTGAAAAAGAAACACAGAACTCGATATATATGAGCCCAATATGAATTTCTT	269
QY	373	gacctgagacaatagctggaatcatctcctcacatttggttccatccatgacaaac	432
DB	268	GCCCTGGAGACAATAGCTGGAATCATTTCTCTACATTTGTTTCATCTGATCAAAAC	209
QY	433	tacatttggttattctcacaataaagtagtgaagtcgttactgcttcttg	492
DB	208	TACATTGTGTTATCTCTCAACAAATAGTCAGTAAAGCTGTACTGCTCTCTG	149
QY	493	ggaatttgattatcatgagactttagaatattgaattatcatctctcgtcttc	552
DB	148	GGAATTTTGATTAATGATGATGACTTTCAGCATTAATTAATTAATTTCTTCTG	89
QY	553	tcaatttggtggtccactcagagagattgattgataaactgttga	603
DB	88	TCAATTTTGGGGTCCACTCAGAGATTGATTTGTAACAATTTTGTGA	38

RESULT 2

LOCUS AA411806 410 bp mRNA EST 12-AUG-1997
 DEFINITION zt67a03.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:727372
 3' similar to SW:CD20_HUMAN P11836 B-LYMPHOCTE ANTIGEN CD20 ;,
 mRNA sequence.
 ACCESSION AA411806 GI:2070377
 VERSION AA411806.1 GI:2070377
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 410)
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
 Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,
 Schellenberg, K., Stepien, M., Tan, F., Theising, B., White, Y., Wyllie
 , T., Waterston, R. and Wilson, R.
 WashU-Merck EST Project 1997
 JOURNAL Unpublished (1997)
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert length: 608 Std Error: 0.00
 Seq primer: 41m13 fwd. ET from Amersham
 High quality sequence stop: 401.
 Location/Qualifiers
 1. 410
 /organism="Homo sapiens"
 /db_xref="GDB:5924284"
 /db_xref="taxon:9606"
 /clone="IMAGE:727372"
 /clone_1ib="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech Laboratories
 , Inc., and primed with a Not I - oligo(dT) primer [5'
 TGTTCACATCTGAAGTGGAGCGGCCGCCCAATTTTCTTTTCTTTT 3'].
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization to Cot5, and was
 constructed by Bento Soares and M. Fatima Bonaldo."
 BASE COUNT 147 a 68 c 76 g 119 t
 ORIGIN
 Query Match 44.3%; Score 267.4; DB 7; Length 410;
 Best Local Similarity 96.1%; Pred. No. 3.6e-58;
 Matches 274; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
 Oy 81 actttcagccagcactttcaactcaagaagccctgcaaaaattttgttagaanaa 140
 Db 410 actttcagccagcactttcaactcaagaagccctgcaaaaattttgttagaanaa 351
 Oy 141 gaaaaacttagagactacacagatcctgttggaaatagacatttcttggagat 200
 Db 350 gaaaaacttagagactacacagatcctgttggaaatagacatttcttggagat 291
 Oy 201 ctctcttccacttggttaaaacataccaggttcccttaatttcttcagagata 260
 Db 290 ctctcttccacttggttaaaacataccaggttcccttaatttcttcagagata 231
 Oy 261 tccatctggggctcgtttgttcaatattcaggaccttccatattgcaatggaag 320
 Db 230 tccatctggggctcgtttgttgttcaatattcaggaccttccatattgcaatggaag 171
 Oy 321 aaaaaccacagaactctgataatattgagccgataatgaatc 365

Db 170 AAAAACCACAGAACTCTGCGAATTTGATTACATTCATGACCTT 126
 RESULT 8
 AA781801/c
 LOCUS
 DEFINITION a158401.s1 Soares_testis_NHT Homo sapiens cDNA clone 1375056 3'
 similar to SW:CD20_HUMAN P11836 B-LYMPHOCTE ANTIGEN CD20 ;, mRNA
 sequence.
 ACCESSION AA781801 GI:2841132
 VERSION AA781801
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 389)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapsb@mail.nih.gov
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
 , Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 www.bio.lnl.gov/dbp/image/image.html
 Insert length: 983 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 311.
 Location/Qualifiers
 1. 389
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="1375056"
 /clone_1ib="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech Laboratories
 , Inc., and primed with a Not I - oligo(dT) primer [5'
 TGTTCACATCTGAAGTGGAGCGGCCGCCCAATTTTCTTTTCTTTT 3'].
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization to Cot5, and was
 constructed by Bento Soares and M. Fatima Bonaldo."
 BASE COUNT 140 a 66 c 68 g 115 t
 ORIGIN
 Query Match 40.7%; Score 245.4; DB 11; Length 389;
 Best Local Similarity 95.8%; Pred. No. 1.5e-52;
 Matches 252; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
 Oy 103 acccaagcccttgcaaaaattatttgcagaaaatgaanaacttagagactacag 162
 Db 389 acccaagcccttgcaaaaattatttgcagaaaatgaanaacttagagactacag 330
 Oy 163 atccgtttggaattagacatttcttggagttacttcccttccacttgtaaaa 222
 Db 329 atccgtttggaattagacatttcttggagttacttcccttccacttgtaaaa 270
 Oy 223 ccaatcagaaggttcccttataattcttcaggagatccatctggggctcgtttg 282
 Db 269 ccaatcagaaggttcccttataattcttgcagatcattcttgaggctctgttttg 210
 Oy 283 ttcattaatcttgagccttccatattgcagtgaaaagaaaacacagaaactcgata 342

Db 209 TTGATTAATTCGTGAGCCCTTCCTAATTCAGTGAAGAAAAACACAGAAACTCTGGA 150
QY 343 atattgagccgaataatcatct 365
Db 149 ATTTGATTACATTGATGACTTT 127

RESULT 9
AA470059 415 bp mRNA EST 09-NOV-1997
LOCUS 2794005.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:730041
DEFINITION 5 Similar to SW:CD20_HUMAN P1836 B-LYMPHOCYTE ANTIGEN CD20 ;
mRNA sequence.
ACCESSION AA470059
VERSION AA470059.1 GI:2197368
KEYWORDS EST.
ORCE human.
ORGANISM Homo sapiens
REFERENCE Homo sapiens; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Maira, M., Martin,
J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
TITLE Unpublished (1997)
JOURNAL Contact: Wilson RK
COMMENT Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 656 Std Error: 0.00
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 393.

FEATURES
source
1. 415
/organism="Homo sapiens"
/db_xref="GDB:5926570"
/db_xref="taxon:9606"
/clone="IMAGE:730041"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5].
TGTTCACATCTGAAGTGGAGCGCGCCCAATTTTCTTTTCTTTT 3']
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 129 a 96 c 67 g 123 t
ORIGIN

Query Match 40.3%; Score 242.8; DB 7; Length 415;
Best Local Similarity 77.3%; Pred. No. 7.1e-52;
Matches 357; Conservative 0; Mismatches 2; Indels 103; Gaps 1;

QY 1 atgattcaagcagcgaacagtcggtgttcgtgtatctccccaagaatcactgct 60
Db 57 ATGCAATTCAGACACCGCACACAGTCCGCTGTTCGTATTTCCTCCAGAAATACACTGCT 116
QY 61 tcgaatatgaatccacaagactttagcgaagactttcaactcaagcccttgcaa 120
Db 117 TCAGAAATATGAGTCCACAGACTTTTCAAGCCAGCACTTTTCAACTCAAAAGCCCTTGCAA 176

QY 121 aaattattgctagaanaaataatcttaggacctatccagatccctgttggaaattag 180
Db 177 AAATTAATTTGCTAGAAAAAATGAATCTTAGG----- 208
QY 181 accttctcttggagttatcttcccttcaacctgttaaaacatatccaaggttccc 240
Db 209 ----- 208
QY 241 ttatatcttcttcagagatatccattctgggctctgtttgttcaatattctggacc 300
Db 209 -----GGATATCCATTCCTGGGCGCTCTGTTCATTAATTCGGAGCC 253
QY 301 ttcccaattgcagtgaaanaaaccacagaactctgataattatgagccgaataatg 360
Db 254 TTCTTAATTCAGAGTAAGAAAAACACAGAAACTCTGATTAATTCAGCCGAATATAG 313
QY 361 aatctcttagtgccttgagagcaatagctggaatcatcttccacatttggttcatc 420
Db 314 AATTTCTTAGTGCCCTCGGAGCAATAGCTGAATCATCTCTCCATTCATTTGCTTCATC 373
QY 421 ctgatacaaaactaactttgtgttattcttcaccaaatagt 462
Db 374 CTAGATCAAAACTACATTTGTGTTCCTACCAAAATAGT 415

RESULT 10
AI002083/c 482 bp mRNA EST 27-AUG-1998
LOCUS ct38b02.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1619019
DEFINITION 3', mRNA sequence.
ACCESSION AI002083
VERSION AI002083.1 GI:3202120
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Homo sapiens; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 482)
NCI-CGAP http://www.ncbi.nlm.nih.gov/cgi/gap
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
TITLE Unpublished (1997)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Contact: Robert Strausberg, Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrp/image/image.html
Insert Length: 598 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 393.

FEATURES
source
1. 482
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1619019"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5].
TGTTCACATCTGAAGTGGAGCGCGCCCAATTTTCTTTTCTTTT 3']
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."

Clones are available from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html. Seq primer: M13-21
Class: BAC ends.

FEATURES

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2379G7"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"

```

```

/Note=Vector; site_1: HindIII; site_2:
HindIII"

```

BASE COUNT	140 a	108 c	115 g	175 t
ORIGIN				

Query Match	19.28;	Score 115.6;	DB 224;	Length 538;
Best Local Similarity	93.18;	Pred No 2	7e-19.	

	seq. local similarity	99.16%	freq. no.	2,781,197					
Matches	121;	Conservative	0;	Mismatches	9;	Indels	0;	Gaps	0;

b
4/4 TGTCTTCTTTCTATTACAGGGAATTTTGATTCATTGATGACTTTCCAGCATTTATGCAATT 211

534 atcattctctgcctttccaatttggggtgcactcagagattgatttgaaca 5933
|||||
212 attcaatttctctgcctttccaatttggggtgcactcagagattgatttgaaca 2712

Y	594	atgttcgtga	603
b	272	ATGTTGTTGA	281

RESULT 15
AA707529/C

LOCUS	AA707529	281 bp	mRNA	EST	12-JAN-1999
DEFINITION	ah1a1a12.s1 Soares_testis_NHT Homo sapiens cDNA clone 1292062 3' , mRNA sequence.				

ACCESSION	AA707529
-----------	----------

VERSION AA707529.1 GI:2717447
KEYWORDS EST.

SOURCE	human
1	human
2	human
3	human
4	human
5	human
6	human
7	human
8	human
9	human
10	human
11	human
12	human
13	human
14	human
15	human
16	human
17	human
18	human
19	human
20	human
21	human
22	human
23	human
24	human
25	human
26	human
27	human
28	human
29	human
30	human
31	human
32	human
33	human
34	human
35	human
36	human
37	human
38	human
39	human
40	human
41	human
42	human
43	human
44	human
45	human
46	human
47	human
48	human
49	human
50	human
51	human
52	human
53	human
54	human
55	human
56	human
57	human
58	human
59	human
60	human
61	human
62	human
63	human
64	human
65	human
66	human
67	human
68	human
69	human
70	human
71	human
72	human
73	human
74	human
75	human
76	human
77	human
78	human
79	human
80	human
81	human
82	human
83	human
84	human
85	human
86	human
87	human
88	human
89	human
90	human
91	human
92	human
93	human
94	human
95	human
96	human
97	human
98	human
99	human
100	human

ORGANISM Homo sapiens

REFERENCE
1 (bases 1 to 281)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

NCI-CCGAP <http://www.nciccgap.org>

TITLE National Cancer Institute, Cancer
Tumor Gene Index

JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

cDNA library arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome
Clone distribution: NCI-CGAP clone distribution
Found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbfp/image/image.html
Insert length: 456 Std Error: 0.00
Seq primer: -40mls fwd. ET from Amersham
High quality sequence stop: 218.

FEATURES

Source

```
/organism="Homo sapiens"  
/db_xref="taxon:9606"  
/clone="1292062"  
/clone_lib="Soares_testis_NHT"  
/sex="male"  
/lab_host="DH10B"
```

/note=Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I. Site_2: Eco RI. 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTGCAACATCTAAGAGGAGCGGGCGCCCAATTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cots, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT	94 a	54 c	45 g	88 t
ORIGIN				

Query Match	18.7%;	Score 113;	DB 11;	Length 281;
Best Local Similarity	100.0%;	Pred. No. 1	1e-18;	
Matches 113;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 491 tgggaatttgatcatgtgacttcagcattatgaattcatctctcgcc 550
|||
Db 148 TGGGAATTTGATTACATTGATGACTTTCAGCATTAATGAATPAATTCATTCTCGCCTT 89

QY 551 ttcacatttcggggtgccaccagagatgtcatgtgacaacatgtttga 603
|||||
Db 88 TCTCAATTTGGGGGTGCCACTCAGAGATTGTGATTGCAACAATGTTGTGA 36

Search completed: October 7, 2001, 03:44:39
Job time: 4344 sec

Thu Oct 11 10:24:05 2001

us-09-735-712-1.rst

Page 12

GenCore version 4.5
Copyright (c) 1993 - 2000, Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 7, 2001, 02:33:15, Search time 1227.86 Seconds
(without alignments)
7596.186 Million cell updates/sec

Title: US-09-735-712-1

Perfect score: 603

Sequence: 1 atgattcaagcaccgcaca.....attgtgacaatgtgttga 603

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba1:*

2: gb_ba2:*

3: gb_ba3:*

4: gb_in1:*

5: gb_in2:*

6: gb_in3:*

7: gb_om:*

8: gb_ov:*

9: gb_pat1:*

10: gb_pat2:*

11: gb_ph:*

12: gb_p11:*

13: gb_p12:*

14: gb_p13:*

15: gb_p14:*

16: em_ba1:*

17: em_ba2:*

18: em_fun:*

19: em_htgo_hum:*

20: em_htgo_inv:*

21: em_htgo_rod:*

22: em_htg_hum1:*

23: em_htg_hum2:*

24: em_htg_hum3:*

25: em_htg_hum4:*

26: em_htg_hum5:*

27: em_htg_hum6:*

28: em_htg_hum7:*

29: em_htg_hum8:*

30: em_htg_inv1:*

31: em_htg_inv2:*

32: em_htg_other:*

33: em_htg_rod:*

34: em_hum1:*

35: em_hum2:*

36: em_hum3:*

37: em_hum4:*

38: em_hum5:*

39: em_hum6:*

40: em_hum7:*

41: em_in:*

42: em_om:*

43: em_or:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	598.2	99.2	691	85	AB013103	AB013103 Homo sapi
2	598.2	99.2	694	89	AF237907	AF237907 Homo sapi
3	596.6	98.9	697	89	AF321127	AF321127 Homo sapi
4	157	26.0	138097	71	AC027787	AC027787 Homo sapi
5	157	26.0	147788	64	AC015840	AC015840 Homo sapi
6	157	26.0	161039	83	AP003127	AP003127 Homo sapi
7	157	26.0	166804	82	AP001034	AP001034 Homo sapi
8	153.4	25.4	138097	71	AC027787	AC027787 Homo sapi

C	9	153.4	25.4	161039	83	AP003127	Homo sapi
	10	115.6	19.2	166804	82	AP001034	Homo sapi
	11	88	14.6	1076	88	AF068288	Homo sapi
	12	86.4	14.3	916	85	AB013102	Homo sapi
	13	86.4	14.3	1619	89	AF237912	Homo sapi
	14	86.4	14.3	1669	9	AR035695	Sequence
	15	86.4	14.3	1669	9	AR083580	Sequence
	16	63.2	10.5	1646	97	HUMIERB	L35848 Homo sapien
	17	63.2	10.5	1661	9	A68627	Sequence 3
	18	63.2	10.5	1661	9	AR082014	Sequence 3
	19	63.2	10.5	1661	9	I25809	Sequence 3
	20	62.4	10.3	1222	7	SSR236932	Sequence 3
	21	58.6	9.7	708	10	AX101306	Sequence
	22	58.6	9.7	1010	94	MUSECERB	J05019 Mouse mast
	23	55.6	9.2	899	94	AF237910	Mus muscu
	24	55.4	9.2	1163	94	AF237915	Mus muscu
	25	55	9.1	1474	10	I07272	Sequence 19
	26	55	9.1	1474	93	HSCD20	X12530 Human mRNA
	27	55	9.1	1476	10	I07277	Sequence 24
	28	55	9.1	1482	91	BC002807	Homo sapi
	29	55	9.1	1597	93	HSCD20A	X07203 Human mRNA
	30	54.8	9.1	518	94	AF280401	Mus muscu
	31	52.6	8.7	862	85	AB013104	Homo sapi
	32	52.6	8.7	930	89	AF286866	Homo sapi
	33	52.6	8.7	1131	89	AF212240	Homo sapi
	34	52.6	8.7	1152	89	AF237908	Homo sapi
	35	52.6	8.7	1289	89	AF142409	Homo sapi
	36	52.6	8.7	1488	9	AX017505	Sequence
	37	52.6	8.7	1679	9	AX017610	Sequence
	38	52	8.6	1182	89	AK000224	Homo sapi
	39	51.8	8.6	1146	97	HUMBILYM	M27394 Human B-Lym
	40	51.4	8.5	1480	94	AF237914	Mus muscu
	41	50.8	8.4	1687	94	AF237909	Mus muscu
	42	49.4	8.2	969	89	AF201951	Homo sapi
	43	49.4	8.2	1012	85	AB026043	Homo sapi
	44	49.4	8.2	1228	89	AF237916	Homo sapi
	45	49.4	8.2	1228	94	AF237913	Mus muscu

ALIGNMENTS

RESULT 1
 LOCUS AB013103 691 bp mRNA PRI 20-MAR-2001
 DEFINITION Homo sapiens mRNA for MS4A5, complete cds.
 ACCESSION AB013103
 KEYWORDS MS4A5; CD20-like 2.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens

REFERENCE 1 (sites)
 Ishibashi, K., Suzuki, M., Sasaki, S. and Imai, M.
 Identification of a new multigene four-transmembrane family (MS4A)
 related to CD20, HTm4 and beta subunit of the high-affinity Ige
 receptor

JOURNAL Gene 264 (1), 87-93 (2001)
 MEDLINE 21142397
 REFERENCE 2 (sites)

AUTHORS Ishibashi, K., Sasaki, S. and Marumo, F.
 TITLE Cloning of three CD20 homolog from human, putative calcium channels
 JOURNAL Unpublished

REFERENCE 3 (bases 1 to 691)
 AUTHORS Ishibashi, K.
 TITLE Direct Submission

JOURNAL Submitted (20-APR-1998) Kenichi Ishibashi, Tokyo Medical and Dental
 University, 2nd Internal Medicine; Yushima 1-5-45, Bunkyo, Tokyo
 113-8519, Japan (E-mail: kishibashi.med2@med.tmd.ac.jp,
 Tel: 81-3-5803-5223, Fax: 81-3-5803-0132)

FEATURES
 source location/Qualifiers
 1..691

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/issue_type="testis"
52..654
/gene="MS4A5"
52..654
/gene="MS4A5"
/codon_start=1
/product="MS4A5"
/protein_id="BAB18739.1"
/db_xref="GI:11559214"
/translation="MDSSTAHSEVFLVPPPEITASVESTELSATPFSQSLQKLF
RKMLIGTILQILFGIMTFSRGVIFLFLKPYPRPFIISGYPGWSVLFINSGL
IAVKRTTEFLILSRIMNPLSALGAIILITFGFIIDONTICGSHONSQCKAVT
VLEGLITLMTFSIIEFLISLPSFSLIGHSEDCDEQC"
BASE COUNT      195 a      142 c      117 g      237 t
ORIGIN

```

```

Query Match      99.2% Score 598.2; DB 85; Length 691;
Best local Similarity 99.5% Pred. No. 1,4e-127;
Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 atggaattcaagcaccgacacacagtcggtgttctglatctccagaatactcgt 60
|||||
DB 52 ATGGAATCAAGCACCACACAGTCGGTGTTCGTGATTTCTCCAGAAATCACTGCT 111

QY 61 tcgaatatagcttcacagacttccagcagccttccactcaagaaccccttgcaa 120
|||||
DB 112 TCGAATAATAGCTTCACAGAACTTCAGCCAGCACTTTCACTCAAGCCCTTGCAA 171

QY 121 aaatatttcctagaanaaaatgaatcttagggactaccagatcctgttggaattatg 180
|||||
DB 172 AAATATTTCCTAGAAAATGAATCTTAGGGACTATCCAGATCCTGTTGGAATTATG 231

QY 181 accttctcttggagttatcttcttccacttgtaaacatataccaaggtttccc 240
|||||
DB 232 ACCTTCTCTTGGAGTTATCTTCTTCACTTTGTAACCAATATCCAGGTTTCCC 291

QY 241 ttatatcttcttcagagatccatctctggtggccttctgttcataatcttcaggacc 300
|||||
DB 292 TTTATATTTCTTCAGGATATCCATCTGGGGCTCTGTGTCATTAATTCGGAGCC 351

QY 301 ttccctaattgcagtgtaaaagaaacacacagaacactgataatattgagccgataatg 360
|||||
DB 352 TTCTATATTCAGTGAGAAAGAAACACAGAAATCTGATTAATATGACCGGATATAG 411

QY 361 aatcttcttagtgcctcgtgagagaataagctggaatcatcttccacattggtttatc 420
|||||
DB 412 AATTTCTTGTAGTGCCCTGGGAGCAATAGCTGAATCATTTCTCCACATTTGGTTTCATC 471

QY 421 ctgagtaaaactacatttctgtgtatctccacaaaatagtgagtgtaagcctgttact 480
|||||
DB 472 CTAGATCAAAACTACATTTGGTTGTTATTCACCAAAATAGTGATTAAGCGCTTACT 531

QY 481 gtccgtgtcttgggaatttgattacatgatgacttcagcatcttattgaattatcatc 540
|||||
DB 532 GTCTGTCTTCTGGGAATTTGATTACATGATGATCTTCAGCANTATGATTAATTATATT 591

QY 541 tctctgcttcttcaatttgggtgctcactcagagagatgtgattgtgaacaatgttgt 600
|||||
DB 592 TCTCTGCTTCTTCATTAATTTGGGTCCTCAGAGAGATGTGATGTGTAACAATGTGTGT 651

QY 601 tga 603
|||||
DB 652 TGA 654

```

RESULT 2
 LOCUS AF237907 694 bp mRNA PRI 17-APR-2001
 DEFINITION Homo sapiens MS4A5 protein mRNA, complete cds.
 ACCESSION AF237907

VERSION AF237907.1 GI:13649400
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Liang, Y. and Tedder, R.F.
TITLE Identification of a CD20-, Fc epsilon R1beta-, and H1m4-related Gene
Family: Sixteen New MS4A Family Members Expressed in Human and
Mouse
JOURNAL Genomics 72 (2), 119-127 (2001)
REFERENCE
AUTHORS Liang, Y. and Tedder, R.F.
TITLE Direct Submission
JOURNAL Submitted (24-FEB-2000) Immunology, Duke Medical Center, Research
Dr., Durham, NC 27710, USA
FEATURES
source
1.694
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS
55..657
/codon_start=1
/product="MS4A5 protein"
/protein_id="AAK37416.1"
/db_xref="GI:13649401"
/translation="MDSRAHSPVLPVPPETIASPYESTELSATTFNSQPLKFLA
RKMKILGTQILFGIMTFSFGVIFLFTLKPRPRPFITLSGIPMGSVLFNSGAFI
LAVKRTTETLILSRIMFLSALGAIAGIILLTFEILLDQNTICGYSQNSQCAV
VFLGLITLMFSLIELFISLPSFSLGSHSDCDEQCC"
BASE COUNT 198 a 143 c 116 g 237 t
ORIGIN

Query Match 99.2%; Score 598.2; DB 89; Length 694;
Best Local Similarity 99.5%; Pred. No. 1.4e-127;
Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 atgattcaagcaccgacagctcgcggtttctgtgattcttcagagaatcagctgc 60
|||||
55 ATGATTCAAGCACCACGACAGCTCGGTGTTCTGTGATTCTCCAGAAATCAGCT 114
|||||
61 tcagaatagctccacagacttccagcagacacttccaactcaaaccccttgcaa 120
|||||
115 TCAGAAATAGCTCCACAGAACTTTCAGCCAGACCTTTCAACTCAAGCCCTTGCAA 174
|||||
121 aaattattgctagaanaaataccttaggactatccagatcctgttggaaatag 180
|||||
175 AAATTATTGCTAGAAAATGAAAATCTTAGGACTATCCAGATCCTGTTGGAAATTATG 234
|||||
181 acccttcttggaattatctccttcaactgtttaaacatatccagaagtttccc 240
|||||
235 ACCTTTCTTTGGAGTATCTTCTTTCCTTTGTTAAACCATATCCAAAGCTTCCC 294
|||||
241 ttataatcttcagagataatcctcctggggtcctgttggtaattcaatctggagcc 300
|||||
295 TTATATTTCTTTCAGGATATCATTCTGGGCTGCTGTTTGTCAATTAATCTGGAGCC 354
|||||
301 ttccataatgagtgagaagaanaaacacagaaactcgtataatattgagcgaaataag 360
|||||
355 TTCCCTAATGCAATGAAAAGAAAACACAGAAACTCTGATATATATGAGCCCAATTAATG 414
|||||
361 aactctcttagtgccctggagagaatagcttggaatcattctcctcaatttggttcacc 420
|||||
415 AATTTTCTTAGTCCCTGGAGCAATAGCTGGAATCAATCTCTCCACATTTGGTTTCAATC 474
|||||
421 ctgagatcaaaactacatttgtgttlaattccacaaaatagtcagtgtaagctgttact 480
|||||
475 CTGAGTCAAAACTACATTGTGTGTTATCTTCACCAAAATAGTCAAGTGAAGCTGTACT 534
|||||
481 gtccctgttctgggaattttgatataatgagacttcaagcaattatgaattatcatt 540
|||||
535 GTCCCTGTCTTGGGAATTTGATTAATGATGACATTTTCAGCATTAATTAATTAATCAAT 594
|||||

QY 541 tctctgcttcttcaatttggggtgcccactcagagagattgtgttgaaacattgt 600
DB 595 TCTCTGCCCTTTCCTCAATTTGGGGTGCACACTCAGAGATTGTGATGTAACATGTTGT 654
QY 601 tga 603
DB 655 TGA 657

RESULT 3
AF321127
LOCUS AF321127 697 bp mRNA PRI 07-FEB-2001
DEFINITION Homo sapiens testis-expressed transmembrane-4 protein (TETM4) mRNA,
complete cds.
ACCESSION AF321127
VERSION AF321127.1 GI:12698681
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Hulet, M.D., Pagler, E., Hornby, J.R., Hogarth, P.M., Eyre, H.J.,
Baker, E., Crawford, J., Sutherland, G.R., Ooms, S.J., and Parish, C.R.
TITLE Isolation, tissue distribution, and chromosomal localization of a
novel testis-specific human four-transmembrane gene related to CD20
and Fc epsilon R1-beta
JOURNAL Biochem. Biophys. Res. Commun. 280 (1), 374-379 (2001)
REFERENCE
MEDLINE 21092614
AUTHORS Hulet, M.D.
TITLE Direct Submission
JOURNAL Submitted (14-NOV-2000) Division of Immunology and Cell Biology,
John Curtin School of Medical Research, Mills Road, Canberra, ACT
2601, Australia
FEATURES
source
1.697
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q12"
1.697
/gene="TETM4"
57..659
/gene="TETM4"
/codon_start=1
/product="testis-expressed transmembrane-4 protein"
/protein_id="AAK01641.1"
/db_xref="GI:12698682"
/translation="MDSRAHSPVLPVPPETIASPYESTELSATTFNSQPLKFLA
RKMKILGTQILFGIMTFSFGVIFLFTLKPRPRPFITLSGIPMGSVLFNSGAFI
LAVKRTTETLILSRIMFLSALGAIAGIILLTFEILLDQNTICGYSQNSQCAV
VLLGLITLMFSLIELFISLPSFSLGSHSDCDEQCC"
BASE COUNT 199 a 142 c 119 g 237 t
ORIGIN

Query Match 98.9%; Score 596.6; DB 89; Length 697;
Best Local Similarity 99.3%; Pred. No. 3.2e-127;
Matches 599; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1 atgattcaagcaccgacagctcgcggtttctgtgattcttcagagaatcagctgc 60
|||||
57 ATGATTCAAGCACCACACACAGTCCGGTGTGTTCTGATTTCTCCAGAAATCAGCTGT 116
|||||
61 tcagaatagctccacagacttccagcagacacttccaactcaaaccccttgcaa 120
|||||
117 TCAGAAATAGCTCCACAGAACTTTCAGCCAGACCTTTCAACTCAAGCCCTTGCAA 176
|||||
121 aaattattgctagaanaaataccttaggactatccagatcctgttggaaatag 180
|||||
177 AAATTATTGCTAGAAAATGAAAATCTTAGGACTATCCAGATCCTGTTGGAAATTATG 236
|||||

[illegible]

Query Match	26.0%;	Score 157;	DB 71;	Length 138097;
Best Local Similarity	91.7%;	Pred. No. 4.8e-26;		
Matches 166;	Conservative 0;	Mismatches 15;	Indels 0;	Gaps 0;

Qy	336	tctgataataattgagccgaataatgtaatcttcttagtgcctgaagagcaataagctggaat	395
Db	102667	TCAGATAAATATGAGCCGGAATATGTAATTTTCTTAGTGCCTGGAGACATAAGCTGGAAT	102608

Qy 396 catctcctacattggttcacatcagatacaaaactacattgtgtattctcacca 455
|||||
Db 102607 CATTCCTCCACATTGGTTTCATCTAGATCAAAACTACATTGTGGTTATTTCACCA 102548

[illegible]

QY	516	t	516
		1	
Db	102487	T	102487

RESULT	5
AC015840	

ACCESSION	DEFINITION	LOCUS	DATE	SIZE	TYPE	ORGANISM
AC015840	homo sapiens clone RP11-24D1, WORKING DRAFT SEQUENCE, 5 unordered pieces.	247100 bp	11/03/90	217	GEN	HOMO SAPIENS

ORGANISM	Homo sapiens
VERSION	ACU13840.2 GI:17031750
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE	human.

REFERENCE
AUTHORS
Birren, B., Linton, L., Nusbaum, C. and Landger, F.
1 (bases 1 to 147788)
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 147788)

TITLE	JOURNAL	REFERENCE
Homo sapiens, clone RP11-24D1	Unpublished	
2 (bases 1 to 147788)		
Birco B	Rintco I	Nichau C
		Randor E
		Allon N
		Andorsee M

Baldwin, J., Barna, N., Beckerly, R., Boguslavskiy, L., Boukhgalter, B., Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Domingo, M., Donelan, L., Doyle, M., Earls, D., Fitzpatrick, W., Forrester, C., Franks, P., Gale, D.,

Galanag, J., Gardyna, S., Grant, G., Hagos, B., Hearford, A., Horton, L.
Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Lehoccky, J., Lieu, C., Locke, K., Macdonald, P., Marguis, N.,
McGuire, D., Neuman, M., O'Connor, T., O'Connell, J., O'Sullivan, T.

Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,

Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome

Direct Submission
Submitted (17-NOV-1999) Whitehead Institute/MIT Center
Research, 320 Charles Street, Cambridge, MA 02141, USA

*	1	42322: contig of 42322 bp in length
*	42323	42422: gap of 100 bp
*	42423	63418: contig of 20996 bp in length
*	63419	63518: gap of 100 bp
*	63519	77966: contig of 14451 bp in length
*	77970	78059: gap of 100 bp
*	78070	93162: contig of 15113 bp in length
*	93183	93282: gap of 100 bp
*	93283	104093: contig of 10811 bp in length
*	104094	104193: gap of 100 bp
*	104194	115702: contig of 11509 bp in length
*	115703	115802: gap of 100 bp
*	115803	125183: contig of 9381 bp in length
*	125184	125283: gap of 100 bp
*	125884	132558: contig of 7275 bp in length
*	132559	132658: gap of 100 bp
*	132659	137219: contig of 4561 bp in length
*	137220	137319: gap of 100 bp
*	137320	143276: contig of 5957 bp in length
*	143277	143376: gap of 100 bp
*	143377	148922: contig of 5546 bp in length
*	148923	149022: gap of 100 bp
*	149023	154159: contig of 5137 bp in length
*	154160	154259: gap of 100 bp
*	154260	157686: contig of 3427 bp in length
*	157687	157786: gap of 100 bp
*	157787	159234: contig of 1448 bp in length
*	159235	159334: gap of 100 bp
*	159335	161039: contig of 1705 bp in length.

```
misc_feature 1. .42322
             /note="assembly-fragment"
misc_feature 42423. 63418
             /note="assembly-fragment"
misc_feature 63319. 77969
             /note="assembly-fragment"
misc_feature 78070. 93182
             /note="assembly-fragment"
misc_feature 93283. 104093
             /note="assembly-fragment"
misc_feature 104194. 115702
             /note="assembly-fragment"
misc_feature 113803. 125183
             /note="assembly-fragment"
misc_feature 125284. 132558
             /note="assembly-fragment"
misc_feature 132659. 137219
             /note="assembly-fragment"
misc_feature 137320. 143276
             /note="assembly-fragment"
misc_feature 143377. 148922
             /note="assembly-fragment"
misc_feature 149023. 154159
             /note="assembly-fragment clone_end:SP6 vector_side:left"
```

Query Match	26.0%;	Score 157;	DB 83;	Length 161039;
Best Local Similarity	91.7%;	Pred. No. 4.9e-26;		
Matches 166;	Conservative	0;	Mismatches 15;	Indels 0;

Qy 396 catctctccacattggtttcatccctagatcaaaactacattggtttatctcacca 455
|||||
Db 45988 CATTCCTCCACATTGGTTTCATCCCTAGATCAAAACTACATTTGGTTATTCACCA 46047
|||||

QY	456	aaatgtagcaggtaaagctgttctcgtcccggtcttctggaaatttggatcatgatgac	515
Db	46048	AAATAGTCAGGTAAAGCTGTACTGTCCGTCTTCTGGTAGTAGTGCATTATAGAG	46107

QY	516	t	516
Db	46108	T	46108

RESULT	7
AP001034/c	
LOCUS	AP001034 166804 bp DNA
DEFINITION	Homo sapiens chromosome 11 clone RP11-729B4 map 11q12, WORKING DRAFT SEQUENCE, 19 unordered pieces.

VERSION AP001034.4 GI:11176992
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens DNA, clone:Rp11-729B4.
ORGANISM Homo sapiens

REFERENCE
1 (bases 1 to 166804)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 166804)

AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Matanabe, H. and Sakaki, Y.
TITLE Homo sapiens 16S, 804 genomic DNA of 11q12
JOURNAL Published only in database (2000) In press
REFERENCE 2 (bases 1 to 166804)
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P.,

TITLE
Fujiiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Submitted (06-JAN-2000) Masahira Hattori, The Institute of Physical
JOURNAL

1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hattori@isc.riken.go.jp, URL: http://hqp.isc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
On Nov 14, 2000 this sequence replaced gi:8117704.

Center: RIKEN Genomic Sciences Center(GSC)

Center code: RIKEN
Web site: <http://b>

Contact: hattori@sc.riken.go.jp

----- Project Information -----
Customer project name: IliadDraft

Center clone name: RP11-729B4

Summary Statistics

Chemistry: Dye-terminator ET-

```
Assembly program: Phrap; version 0.990329
```

Consensus quality: 162254 bases at least Q40

consensus quality: 1b3858 bases at least Q30

Consensus quality: 164544 bases at least Q20
Insert size: 165004; sum-of-continigs
Quality coverage: 9.28x in Q20 bases; sum-of-continigs

NOTE: This is a 'working draft' sequence. It currently consists of 19 continigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the continigs are represented as runs 'N', but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1 22398 contig of 22398 bp in length
22499 45238 contig of 22740 bp in length
45339 64278 contig of 18940 bp in length
64379 77579 contig of 13201 bp in length
77680 93718 contig of 16039 bp in length
93819 108057 contig of 14239 bp in length
108158 117006 contig of 8849 bp in length
117107 126302 contig of 9196 bp in length
126403 133479 contig of 7077 bp in length
133580 140425 contig of 6846 bp in length
140526 144653 contig of 4128 bp in length
144754 149354 contig of 4601 bp in length
149455 154522 contig of 5068 bp in length
154623 155819 contig of 1297 bp in length
156020 159504 contig of 3485 bp in length
159605 159679 contig of 75 bp in length
159780 162418 contig of 2639 bp in length
162519 165383 contig of 2865 bp in length
165484 166804 contig of 1321 bp in length.

NOTE: This is a 'working draft' sequence. It currently consists of 19 continigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the continigs are represented as runs of 'N', but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 22398: contig of 22398 bp in length
22399 22498: gap of 100 bp
22499 45238: contig of 22740 bp in length
45239 45338: gap of 100 bp
45339 64278: contig of 18940 bp in length
64279 64378: gap of 100 bp
64379 77579: contig of 13201 bp in length
77580 77679: gap of 100 bp
77680 93718: contig of 16039 bp in length
93719 93818: gap of 100 bp
93819 108057: contig of 14239 bp in length
108058 108157: gap of 100 bp
108158 117006: contig of 8849 bp in length
117007 117106: gap of 100 bp
117107 126302: contig of 9196 bp in length
126303 126402: gap of 100 bp
126403 133479: contig of 7077 bp in length
133480 133579: gap of 100 bp
133580 140425: contig of 6846 bp in length
140426 140525: gap of 100 bp
140526 144653: contig of 4128 bp in length
144654 144753: gap of 100 bp
144754 149354: contig of 4601 bp in length
149355 149454: gap of 100 bp
149455 154522: contig of 5068 bp in length
154523 154622: gap of 100 bp
154623 155819: contig of 1297 bp in length
155820 156019: gap of 100 bp
156020 159504: contig of 3485 bp in length
159505 159604: gap of 100 bp
159605 159679: contig of 75 bp in length
159680 159779: gap of 100 bp
159780 162418: contig of 2639 bp in length
162419 162518: gap of 100 bp
162519 165383: contig of 2865 bp in length

* 165384 165483: gap of 100 bp
* 165484 166804: contig of 1321 bp in length.
FEATURES
Location/Qualifiers
1.166804
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q12"
/clone="RP11-729B4"

misc.feature 1.22398
/note="assembly-fragment"
misc.feature 22499..45238
/note="assembly-fragment"
misc.feature 45339..64278
/note="assembly-fragment"
misc.feature 64379..77579
/note="assembly-fragment"
misc.feature 77680..93718
/note="assembly-fragment"
misc.feature 93819..108057
/note="assembly-fragment"
misc.feature 108158..117006
/note="assembly-fragment"
misc.feature 117107..126302
/note="assembly-fragment"
misc.feature 126403..133479
/note="assembly-fragment"
misc.feature 133580..140425
/note="assembly-fragment"
misc.feature 140526..144653
/note="assembly-fragment"
misc.feature 144754..149354
/note="assembly-fragment"
misc.feature 149455..154522
/note="assembly-fragment"
misc.feature 154623..155819
/note="assembly-fragment"
misc.feature 156020..159504
/note="assembly-fragment"
misc.feature 159605..159679
/note="assembly-fragment clone_end:17 vector_side:left"
misc.feature 159780..162418
/note="assembly-fragment"
misc.feature 162519..165383
/note="assembly-fragment"
misc.feature 165484..166804
/note="assembly-fragment"

BASE COUNT 50706 a 31523 c 32236 g 50478 t 1801 others
ORIGIN

Query Match 26.0%; Score 157; DB 82; Length 166804;
Best Local Similarity 91.7%; Pred. No. 4.9e-26;
Matches 166; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 336 tctgataatattgagccgaataatgaattctcttagtccttgagagcaatgctgnaat 395
|||
DB 7669 TCAGATAAATATGAGCGAATATGAATTTCTTAGGCCCTGGAGCAATACCTGAAT 7610
|||
QY 396 catctccctcaatttggttcatctctagatcaaaactacatttggttattctacca 455
|||
DB 7609 CATTCCCTCCACATTGGTTCACTCCTAGATCAAACTACATTGGTTATTCACCA 7550
|||
QY 456 aatatgcaagtgaagcgtctctctcgtctcgttggaatttatacatatgtagc 515
|||
DB 7549 AAATGACAGTGAAGCTGTACTGCTGTCTTGTAATATGTGCATTATAGAG 7490
|||
QY 516 t 516
|
DB 7489 T 7489

RESULT 8

AC027787
LOCUS AC027787 138097 bp DNA HTG 18-AUG-2000
DEFINITION Homo sapiens chromosome 15 clone RP11-196E16 map 15, LOW-PASS
SEQUENCE SAMPLING.
AC027787
VERSION AC027787.2 GI:9845160
KEYWORDS HTG; HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Birren, B., Linton, L., Nusbaum, C., and Lander, E.
TITLE 1 (bases 1 to 138097)
JOURNAL Homo sapiens chromosome 15, clone RP11-196E16
REFERENCE
AUTHORS unpublished
2 (bases 1 to 138097)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boguslavsky, L., Bouckhalter, B., Brown, A., Burkett, G.,
Campolino, A., Castle, A., Choquel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., DeRellano, K., Dewar, K., Diaz, J.S.,
Dodgson, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Larocque, K., Lamazares, R., Landers, T., Lehotzky, J.,
Levin, R., Lien, C., Liu, G., Locke, K., MacDonald, P., Marquis, N.,
McCarthy, M., McKean, P., McGurk, A., McKernan, K., McPheeters, R.,
Meldrum, J., Meneses, L., Miho, T., Miranda, C., Mlenka, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange, R., Stojanovic, N., Subramanian, A., Talamas, J.,
Teste, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (01-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 18, 2000 this sequence version replaced gi:7382629.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996.1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L9908
Center clone name: 196_E_16

* NOTE: This record contains 163 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1
* 664 763: gap of 663 bp in length
* 764 1464: contig of 701 bp in length
* 1465 1564: gap of 100 bp
* 1565 2268: contig of 704 bp in length
* 2269 2368: gap of 100 bp
* 2369 3042: contig of 674 bp in length
* 3043 3142: gap of 100 bp
* 3143 3820: contig of 678 bp in length

* 3821 3920: gap of 100 bp
* 3921 4626: contig of 706 bp in length
* 4627 4726: gap of 100 bp
* 4727 5423: contig of 697 bp in length
* 5424 5523: gap of 100 bp
* 5524 6234: contig of 711 bp in length
* 6235 6334: gap of 100 bp
* 6335 7040: contig of 706 bp in length
* 7041 7140: gap of 100 bp
* 7141 7837: contig of 697 bp in length
* 7838 7937: gap of 100 bp
* 7938 8632: contig of 695 bp in length
* 8633 8732: gap of 100 bp
* 8733 9430: contig of 698 bp in length
* 9431 9530: gap of 100 bp
* 9531 10214: contig of 684 bp in length
* 10215 10314: gap of 100 bp
* 10315 10980: contig of 666 bp in length
* 10981 11080: gap of 100 bp
* 11081 11783: contig of 703 bp in length
* 11784 11883: gap of 100 bp
* 11884 12582: contig of 699 bp in length
* 12583 13389: contig of 707 bp in length
* 13390 13489: gap of 100 bp
* 13490 14186: contig of 697 bp in length
* 14187 14286: gap of 100 bp
* 14287 14996: contig of 710 bp in length
* 14997 15096: gap of 100 bp
* 15097 15796: contig of 700 bp in length
* 15797 15896: gap of 100 bp
* 15897 16583: contig of 687 bp in length
* 16584 16683: gap of 100 bp
* 16684 17376: contig of 693 bp in length
* 17377 17476: gap of 100 bp
* 17477 18178: contig of 702 bp in length
* 18179 18278: gap of 100 bp
* 18279 18962: contig of 684 bp in length
* 18963 19062: gap of 100 bp
* 19063 19673: contig of 611 bp in length
* 19674 19773: gap of 100 bp
* 19774 20436: contig of 663 bp in length
* 20437 20536: gap of 100 bp
* 20537 21233: contig of 697 bp in length
* 21234 21333: gap of 100 bp
* 21334 22042: contig of 709 bp in length
* 22043 22142: gap of 100 bp
* 22143 22808: contig of 666 bp in length
* 22809 22908: gap of 100 bp
* 22909 23595: contig of 687 bp in length
* 23596 23695: gap of 100 bp
* 23696 24391: contig of 696 bp in length
* 24392 24491: gap of 100 bp
* 24492 25190: contig of 699 bp in length
* 25191 25290: gap of 100 bp
* 25291 25981: contig of 691 bp in length
* 25982 26081: gap of 100 bp
* 26082 26765: contig of 684 bp in length
* 26766 26865: gap of 100 bp
* 26866 27563: contig of 698 bp in length
* 27564 27663: gap of 100 bp
* 27664 28377: contig of 714 bp in length
* 28378 28477: gap of 100 bp
* 28478 29184: contig of 707 bp in length
* 29185 29284: gap of 100 bp
* 29285 29983: contig of 699 bp in length
* 29984 30083: gap of 100 bp
* 30084 30775: contig of 692 bp in length
* 30776 30875: gap of 100 bp
* 30876 31565: contig of 690 bp in length
* 31566 31665: gap of 100 bp
* 31666 32360: contig of 695 bp in length
* 32361 32460: gap of 100 bp

```

* 32461 33157: contig of 697 bp in length
* 33158 33257: gap of 100 bp
* 33258 33951: contig of 694 bp in length
* 33952 34051: gap of 100 bp
* 34052 34735: contig of 684 bp in length
* 34736 34835: gap of 100 bp
* 34836 35483: contig of 648 bp in length
* 35484 35583: gap of 100 bp
* 35584 36308: contig of 725 bp in length
* 36309 36408: gap of 100 bp
* 36409 37108: contig of 700 bp in length
* 37109 37208: gap of 100 bp
* 37209 37892: contig of 684 bp in length
* 37893 37992: gap of 100 bp
* 37993 38672: contig of 680 bp in length
* 38673 38772: gap of 100 bp
* 38773 39475: contig of 703 bp in length
* 39476 39575: gap of 100 bp
* 39576 40290: contig of 715 bp in length
* 40291 40390: gap of 100 bp
* 40391 41087: contig of 697 bp in length
* 41088 41187: gap of 100 bp
* 41188 41891: contig of 704 bp in length
* 41892 41991: gap of 100 bp
* 41992 42679: contig of 688 bp in length
* 42680 42779: gap of 100 bp
* 42780 43500: contig of 721 bp in length
* 43501 43600: gap of 100 bp
* 43601 44295: contig of 695 bp in length
* 44296 44395: gap of 100 bp
* 44396 45088: contig of 693 bp in length
* 45089 45188: gap of 100 bp
* 45189 45880: contig of 692 bp in length
* 45881 45980: gap of 100 bp
* 45981 46683: contig of 703 bp in length
* 46684 46783: gap of 100 bp
* 46784 47481: contig of 698 bp in length
* 47482 47581: gap of 100 bp
* 47582 48290: contig of 709 bp in length
* 48291 48390: gap of 100 bp
* 48391 49077: contig of 687 bp in length
* 49078 49177: gap of 100 bp
* 49178 49879: contig of 702 bp in length
* 49880 49979: gap of 100 bp
* 49980 50684: contig of 705 bp in length
* 50685 50784: gap of 100 bp
* 50785 51507: contig of 723 bp in length
* 51508 51607: gap of 100 bp
* 51608 52320: contig of 713 bp in length
* 52321 52420: gap of 100 bp
* 52421 53133: contig of 713 bp in length
* 53134 53233: gap of 100 bp
* 53234 53933: contig of 700 bp in length
* 53934 54033: gap of 100 bp
* 54034 54703: contig of 670 bp in length

```

Query Match 25.4%; Score 153.4; DB 71; Length 138097;
 Best Local Similarity 96.3%; Pred. No. 3.2e-25;
 Matches 157; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 1 atgagattcaagcagcagcagcagtcggtttcttgatattctcctcagaataactgct 60
    |||||||
DB 109596 ATGGATTCAAGCAGCAGCAGCAGTCCGTTCTGTTCTGTTCTCCAGAAATCAGTGT 109655
    |||||||

QY 61 tcgaataatagtcacagacacttcagcagcagccttcaactcaaaccccttgcaa 120
    |||||||
DB 109656 TCAGAATATAGTCCAGACACTTCAGCCACGACCTTTTCAACTCAAGCCCTTGCAA 109715
    |||||||

QY 121 aaattatttctagaanaaatgaaatcttagagactatccaga 163
    |||||||
DB 109716 AAATTTATTCTGTAAGAAATGAAATCTTAGGGTAACTAGA 109758
    |||||||

```

```

RESULT 9
AP003127/c
LOCUS
DEFINITION
Homo sapiens chromosome 11 clone RP11-71066 map 11q, WORKING DRAFT
SEQUENCE, 15 unordered pieces.
AP003127
ACCESSION
AP003127.1 GI:12597183
VERSION
HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 161039)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 161,039 genomic DNA of 11q
Published only in Database (2001) In press
2 (bases 1 to 161039)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (26-JAN-2001) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsr.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsr.riken.go.jp
----- Project Information
Center project name: HumDrat11
Center clone name: RP11-71066
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 157456 bases at least Q40
Consensus quality: 158960 bases at least Q30
Consensus quality: 159457 bases at least Q20
Insert size: 159639; sum-of-contigs
Quality coverage: 8.61x in Q20 bases; sum-of-contigs

```

NOTE: This is a 'working draft' sequence. It currently consists of 15 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs 'N', but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```

1 42322 contig of 42322 bp in length
42423 63418 contig of 20996 bp in length
63519 77969 contig of 14451 bp in length
78070 93182 contig of 15113 bp in length
93283 104093 contig of 10811 bp in length
104194 115702 contig of 11509 bp in length
115803 125183 contig of 9581 bp in length
125284 132558 contig of 7275 bp in length
132659 137219 contig of 4561 bp in length
137320 143276 contig of 5957 bp in length
143377 148922 contig of 5546 bp in length
149023 154159 contig of 5137 bp in length
154260 157686 contig of 3427 bp in length
157787 159234 contig of 1448 bp in length
159335 161039 contig of 1705 bp in length.

```

* NOTE: This is a 'working draft' sequence. It currently consists of 15 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of 'N', but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will

```

* be preserved.
* 1 42322: contig of 42322 bp in length
* 42323 42422: gap of 100 bp
* 42423 63418: contig of 20996 bp in length
* 63419 63518: gap of 100 bp
* 63519 77969: contig of 14451 bp in length
* 77970 78069: gap of 100 bp
* 78070 93182: contig of 15113 bp in length
* 93183 93282: gap of 100 bp
* 93283 104093: contig of 10811 bp in length
* 104094 104193: gap of 100 bp
* 104194 115702: contig of 11509 bp in length
* 115703 115802: gap of 100 bp
* 115803 125183: contig of 9381 bp in length
* 125184 125283: gap of 100 bp
* 125284 132358: contig of 7275 bp in length
* 132359 132658: gap of 100 bp
* 132659 137219: contig of 4561 bp in length
* 137220 137319: gap of 100 bp
* 137320 143276: contig of 5957 bp in length
* 143277 143376: gap of 100 bp
* 143377 148922: contig of 5546 bp in length
* 148923 149022: gap of 100 bp
* 149023 154159: contig of 5137 bp in length
* 154160 154259: gap of 100 bp
* 154260 157686: contig of 3427 bp in length
* 157687 15786: gap of 100 bp
* 15787 159234: contig of 1448 bp in length
* 159235 159334: gap of 100 bp
* 159335 161039: contig of 1705 bp in length.
* Location/Qualifiers
* 1.161039
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /chromosome="11"
  /map="11q"
  /clone="RP11-71066"
  1.42322
    misc_feature
      /note="assembly_fragment"
      42423..63418
    misc_feature
      /note="assembly_fragment"
      63519..77969
    misc_feature
      /note="assembly_fragment"
      78070..93182
    misc_feature
      /note="assembly_fragment"
      93283..104093
    misc_feature
      /note="assembly_fragment"
      104194..115702
    misc_feature
      /note="assembly_fragment"
      115803..125183
    misc_feature
      /note="assembly_fragment"
      125284..132358
    misc_feature
      /note="assembly_fragment"
      132659..137219
    misc_feature
      /note="assembly_fragment clone_end:SP6 vector_side:left"
      137320..143276
    misc_feature
      /note="assembly_fragment"
      143377..148922
    misc_feature
      /note="assembly_fragment clone_end:T7 vector_side:left"
      149023..154159
    misc_feature
      /note="assembly_fragment"
      154260..157686
    misc_feature
      /note="assembly_fragment"
      157787..159234
    misc_feature
      /note="assembly_fragment"
      159335..161039
    misc_feature
      /note="assembly_fragment"
      161039..161039
  misc_feature
    /note="assembly_fragment"
    31166 c 31523 g 47795 t 1400 others
  BASE COUNT
  49155 a 31166 c 31523 g 47795 t 1400 others
  ORIGIN

```

```

Query Match 25.4%; Score 153.4; DB 83; Length 161039;
Best Local Similarity 96.3%; Pred. No. 3.3e-25;

```

```

Matches 157; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 atgattcaagcaccgacacagtcggtttcttgatcttcctcagaatracgct 60
|||||
Db 149618 ATGATTCAAGCACCACGACACAGTCGGTCTTCTGGATTTCCTCAGAAATCAGTGT 149559
|||||
QY 61 tcaaatatgacccacagacttcagcagacacttcacacttcacaaagcccttgca 120
|||||
Db 149558 TCAGAAATATGATGACACAGACTTTCAGCCAGACCTTTCAACTCAAGCCCTTCCAA 149499
|||||
QY 121 aaattattgtagaaaatcgaatccttagggactaccga 163
|||||
Db 149498 AAATTATTGCTAGAAAATGAAATCTTAGGGGTAGTAGA 149456
|||||

RESULT 10
AP001034 166804 bp DNA HTG 14-NOV-2000
LOCUS Homo sapiens chromosome 11 clone RP11-729B4 map 11q12, WORKING
DEFINITION DRAFT SEQUENCE. 19 unordered pieces.
ACCESSION AP001034 GI:11176992
VERSION AP001034.4
KEYWORDS HTG; HTGS; PHASE1; HTGS; DRAFT.
SOURCE Homo sapiens DNA, clone:RP11-729B4.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 166804)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 166,804 genomic DNA of 11q12.
Published Only in Database (2000) In press
2 (bases 1 to 166804)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (06-JAN-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@psc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
On Nov 14, 2000 this sequence version replaced gi:8117704.

COMMENT
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@psc.riken.go.jp
----- Project Information
Center Project name: HumDraft11
Center clone name: RP11-729B4
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 162254 bases at least Q40
Consensus quality: 163858 bases at least Q30
Consensus quality: 164544 bases at least Q20
Insert size: 165004; sum-of-ctrls
Quality coverage: 9.28x in Q20 bases; sum-of-ctrls
-----
NOTE: This is a 'working draft' sequence. It currently consists of
19 contigs. The true order of the pieces is not known and thelir
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved
1 22398 contig of 22398 bp in length
22499 45238 contig of 22740 bp in length
45339 64278 contig of 18940 bp in length
64379 77579 contig of 13201 bp in length
77680 93718 contig of 16039 bp in length
93819 108057 contig of 14239 bp in length

```


ORGANISM	Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.	
REFERENCE	1 (bases 1 to 1619)
AUTHORS	Liang,Y. and Tedder,T.F.
TITLE	Identification of a CD20-, FcepsilonR1beta-, and HtmD-Related Gene Family: sixteen New MSA Family Members Expressed in Human and Mouse
JOURNAL	Genomics 72 (2), 119-127 (2001)
REFERENCE	2 (bases 1 to 1619)
AUTHORS	Liang,Y. and Tedder,T.F.
TITLE	Direct Submission
JOURNAL	Submitted (24-FEB-2000) Immunology, Duke Medical Center, Research Dr., Durham, NC 27710, USA
FEATURES	Location/Qualifiers
Source	1..1619
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/tissue_type="Placenta"
CDS	144..806
	/codon_start=1
	/product="MSA4A protein"
	/protein_id="AAK37594.1"
	/db_xref="GI:13655457"
	/translation="MTTMOGMEQAMPAGGVDPOLGNMAVIHSHTNKLOEFLKGP KVLDVQLITLMSLTGMTCMASNTYGSNPISVYIGTYIMGSVMFTLSGISIAA GIPTTGKIVRSGLNMTSSVLASGILIMPESIAFYSPPHPXCNYYGNSNNCGTMS ILMDLDGMVLLSVLEFCINVSLSAPGCNVLCCTPRGVLLIPSHMAYETASTPLN EV"
BASE COUNT	496 a 320 c 304 g 499 t
ORIGIN	
Query Match	14.3%; Score 86.4; DB 89; Length 1619;
Best Local Similarity	51.1%; Pred. No. 5.1e-10;
Matches 232; Conservative	0; Mismatches 216; Indels 6; Gaps 1;
Oy	119 aaaatatttgctagaaanaatgaatctaaggactacaagaacctggttgaaatta 178
Db	253 AGAATCTTTGTGAAGGAGAACCCAAAGCCTTGGGGTTGTGCAGATTCTGACGCCCTGA 312
Oy	179 tgaccttcctttggagttaattccctcttcaaccttgttaaaccatatcaagtttc 238
Db	313 TGAGCCTTAGCATGGGAAATAACATGATGTATGCACTTAATCTTAATGAGAATAACC 372
Oy	239 cccttatattccttcagatatcatctctcgtggcgctgtttgttccaattaattctgag 298
Db	373 CTATTTCCGTGTAATMGCGGTACACAATTTGGGGGTCAAGTAATGTTATTTTCAGAT 432
Oy	299 cctctctaatggcagtgaaaagaanaaaccaagaagaacctgatataattatggccogaataa 358
Db	433 CCTTTCATTAATGACACAGAAATTAACATACAAAAGSCCTGGTCCGAGGTAGTACAGAA 492
Oy	359 tgaactctcttagtgccctcggagagaactagctggaatcatctcctccacatttggttcca 418
Db	493 TGAATATCACCAAGCTCTGTACTGCTGCATCAGGAGATCTTAATCAACACATTTAGCTTGG 552
Oy	419 tectatgat-----caaactacattgttgattatctccaccaaaatagtcagtgaag 472
Db	553 CGTTTATTCATTTCCATACACCTTACTCTACTAATCTATGAGCAACTCAATATATGTCATG 612
Oy	473 ctgtatgcctcgtctctcggaaatttgattacaatgatgaacttaagcattatgat 532
Db	613 GGACATATGTCATCTTAATGGGTCTGGAATGGCATGTCCTCTCTTAAGTGTGCTGGAAAT 672
Oy	533 tatcaattctctcgtccttctccaatttgggggtg 566
Db	673 TCTGCATTCGTGTGTCCTCTCTGCTGCTTGGATG 706

DEFINITION	Sequence 2 from patent US 5871930.
ACCESSION	AR035695
VERSION	AR035695.1
KEYWORDS	GI:5952363
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 1669)
TITLE	Bandman,O., Lal,P. and Corley,N.C.
JOURNAL	High affinity immunoglobulin E receptor-like protein
FEATURES	Patent: US 5871930-A 2 16-FEB-1999;
source	Location/Qualifiers
	1..1669
BASE COUNT	481 a 350 c 334 g 501 t 3 others
ORIGIN	/organism="unknown"
Query Match	14.3%; Score 86.4; DB 9; Length 1669;
Best local similarity	51.1%; Pred. No. 5.1e-10;
Matches 232;	Conservative 0; Mismatches 216; Indels 6; Gaps
QY	119 aaaaattatctgtagaaaatgaaatcttaaggactatccagatccctgttggaa17
Db	325 AGAATCTCTTGAAGGAGAACCCAAAGCTCTGGGGTGTGCATATCTGATGCCCGA 384
QY	179 tgaccttcttttgagttatcttcctttccactctgttaaaacatccaaagt123
Db	385 TGAGCCTTAGCATGGGAAATACATGATGTGTATGGCATCTTAATCTTATGGAATACC 444
QY	239 ccttataattcttccagaatatccatctctgggctcgttttltcaatattctgag 298
Db	445 CTATTCCGTGATATTCGGGTACACAAATTTGGGGTGAGTATCTTATTTTCAGGAT 504
QY	299 ccttcttaattggcagtgaagaagaacacagaactctgatatattggccgataa 355
Db	505 CTTTCTCAATTTGCACACGAAATTTGAACCTACAAAAGCCTCTGTCGAGTAGTATAGAA 564
QY	359 tgaacttctttagtgccttcgaggaacatagcttgaaatcatctccacatttg141
Db	565 TGAATATACACAGCTCTGTACTGGGTGCATCAGGAGATCTTAATCACACATTAGCTTGG 622
QY	419 tccctgat-----caaaactacatttgg1attctccacaaaatagctagt1aag 477
Db	625 CGTTTATTCATTCATCCATACCCCTTACTCTAATCTAGTATGGAACATCAATTAATGTCATG 684
QY	473 ctgttaactgcctgtctcttggaatttgattcaatgatgaacttcagcat1atgaat 532
Db	685 GGACGATATGCCATCTTAATGGGTCTGAGATGGCAGATGCTCTCTTAAGTGTGCTGGAAT 744
QY	533 tattcaattctctgccttcttcaatttgg9gtg 566
Db	745 TCTGCAATTCGTGTGTCCCTCTCTGCTTGGATG 778
RESULT 15	
LOCUS	AR083580 1669 bp DNA PAT 01-SEP-2000
DEFINITION	Sequence 2 from patent US 5977072.
ACCESSION	AR083580
VERSION	AR083580.1
KEYWORDS	GI:10010353
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 1669)
TITLE	Bandman,O., Lal,P. and Corley,N.C.
JOURNAL	High affinity immunoglobulin E receptor-like protein
FEATURES	Patent: US 5977072-A 2 02-NOV-1999;
source	Location/Qualifiers
	1..1669
BASE COUNT	481 a 350 c 334 g 501 t 3 others
	/organism="unknown"

ORIGIN

Query Match 14.3%; Score 86.4; DB 9; Length 1669;
Best Local Similarity 51.1%; Pred. No. 5.1e-10;
Matches 232; Conservative 0; Mismatches 216; Indels 6; Gaps 1;

```
QY 119 aaaaattatttctagaanaaaatgaaaatcttagagactatccagatcctgttgaatta 178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 325 AGAAGTCTTGAGAGGAGAGAACCCAAAGTCTGTGGGTTGTGACATCTGACTGCCCTGA 384
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 179 tgaccttctcttgaggtatcttcccttccacctgttaaaacatatccaagtttc 238
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 385 TGACCCCTTACATGGGAAATGAATGATGTATGGCATCTAATACTTATGGAAGTAACC 444
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 239 ccttataattcttcaaggtatccatctctgggctctgttcttcaatattcttgag 298
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 445 CTATTTCCGTATATGCGGTACACAATTTGGGGGTGAGTAATGTATTATTTCAGGAT 504
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 299 ccttccataatgtagtaaaagaaacacagaactctgataataatgaagccgaataa 358
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
505 CCTTGTCAATTTGAGAGGAGAAATTAGAACTACAAATTTGGGGGTGAGTAATGTATTTCAGGAA 564
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 359 tgaactcttctagtgccctgaaggaataagctggaatcaltctcctcacatttggttca 418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 565 TGAATATCACACAGCTCTGTACTGCTGCATCAGGATCTTAATCAACACATTAGCTTGG 624
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 419 tccctagat-----caaaactacatttggttattctcaccaaaatagtcagtgaag 472
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 625 CGTTTATTCATTCACATCCCTTACTGTACTGTACTGTATGGCAACTCAAAATTAATGTGATG 684
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 473 ctgttactgcctgcttcttggaatttgattacattgacttgcagcatattgaat 532
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 685 GGACTATGTCCATCTATATGGGCTGTGATGCAATGATGCTCTCTTAAGTGTGCTGGAAT 744
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 533 tattcaattctctgccttctcacaatttggggtg 566
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 745 TCTGCATTGCTGTGCTCCCTCTGCTGCTTTGGATG 778
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: October 7, 2001, 04:07:37
Job time: 5662 sec

THIS PAGE BLANK (USPTO)


```

PS      Claim 1: SEQ ID 3078; 71bp + CD-ROM; English.
XX
CC      The present sequence is one of a large number of 5' ESTs derived from
CC      mRNAs encoding secreted proteins. An ORF has been identified within the
CC      sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC      derived from 30 different tissues. EST sequences usually correspond
CC      mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC      often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC      well suited for isolating cDNA sequences derived from the 5' ends of
CC      mRNAs and even in those cases where longer cDNA sequences have been
CC      obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC      mRNAs with intact 5' ends and can therefore be used to obtain full length
CC      cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC      gene therapy and chromosome mapping procedures. They are used to obtain
CC      upstream regulatory sequences and to design expression and secretion
CC      vectors.
XX
XX      Sequence 372 BP; 105 A; 87 C; 53 G; 117 T; 10 other:
XX
XX      Query Match          46.6%; Score 280.8; DB 21; Length 372;
XX      Best Local Similarity 96.3%; Pred. No. 3e-66;
XX      Matches 289; Conservative 6; Mismatches 4; Indels 1; Gaps 1
XX
OY      1 atgagtcgaagcaccgcacacagtcgcggtgtttctgtatcttcctcagaatacactgct 60
Db      73 atgagtcgaagcaccgcacacagtcgcggtgtttctgtatcttcctcagaatacactgct 132
OY      61 tcagatatgagtcacagaactttcagccacagcacttttaactcaaacgcccttgc-a 119
Db      133 tcaaratatgagtcacagaactttcagccacagcacttttaactcaaacgcccttgc-a 192
OY      120 aaatatttgcctagaaatgaaaaatctttaggaactatccagatcctgttggaaatlat 179
Db      193 aaatatttgcctagaaatgaaaaatctttaggaactatccagatcctgttggaaatlat 252
OY      180 gaccctttcttttgagatcatcttcctttaccctcgttaaaaacatatccaagttcc 239
Db      253 gaccctttcttttgagatcatcttcctttaccctcgttindaacacatatccaagttcc 312
OY      240 ctttatattcttcaagatatccattcggggcctcgtttgttcaatattcggagc 299
Db      313 ctttatattcttcaanaratatccattcggggcctcgtttgttcaatattcggagc 372
XX
RESULT 2
XX      AAF58252
XX      AAF58252 standard; DNA; 936 BP.
XX
AC      AAF58252;
XX
DT      24-APR-2001 (first entry)
XX
DE      Oligonucleotide D1835.
XX
KW      Electron-transfer group; EFM; mismatch; genotyping;
XX      gene expression; ss.
XX
OS      Synthetic.
XX
PN      WO200107665-A2.
XX
PD      01-FEB-2001.
XX
PF      26-JUL-2000; 2000WO-US20476.
XX
PR      26-JUL-1999; 99US-0145695.
XX      17-MAR-2000; 2000US-0190259.
XX
PA      (CLIN-) CLINICAL MICRO SENSORS INC.
XX
XT      Unbek RM;
XX

```

```

DR   WP1, 2001-159728/16.
XX
PT   Nucleic acids containing electron-transfer group, useful as labels in
PT   hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT   a single surface.
XX
PS   Example 6; Page 127; 159pp; English.
XX
CC   The present invention relates to a composition comprising two nucleic
CC   acids each containing an electron transfer group (ETM) having
CC   different redox potentials. The invention is used for electronic
CC   detection of nucleic acids, especially of substitutions (mismatches)
CC   and single-nucleotide polymorphisms, e.g. for genotyping,
CC   monitoring gene expression.
XX
SQ   Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;

Query Match          15.6%; Score 94.2; DB 22; Length 936;
Best Local Similarity 1.1%; Pred. No. 5.2e-16;
Matches 6; Conservative 356; Mismatches 209; Indels 0; Gaps 0;

QY   31 ttctcgtatcttcctccagaatcacctcgtctcaagaatcgtgacacgaacttcagcc 90
DB   151 www.wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 210
    :: : :::: : : :::: : : : :::: : : : : : : : : : : : : : : : :
QY   91 acgaccttcaaccocaaagcccttcgcaaaattatttgcctagaanaatgaaatctta 150
DB   211 www.wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 270
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   151 gggagctatccagatccctgttggaattabgacctttcttgtagtattcctcttc 210
DB   271 www.wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 330
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   211 acctgttaaaccatccaaagttcccttatattcttcaagatccattctgg 270
DB   331 www.wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 390
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   271 ggcctcgttttgcattaatcttgagcccttcctaattgacgtgaaagaaacacaca 330
DB   391 wgcttawwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 450
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   331 gaacctcgtataatttgagccgaataatgaattcttctagtgcctgagacataagct 390
DB   451 www.wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 510
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   391 ggaatcatctcctacatttggttccatcctcagaacaaactacattgtgtattct 450
DB   511 www.wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 570
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   451 caaccaaatagtcagtgaagcgttactcgtcctgcttccttggaatttgatacatg 510
DB   571 www.wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 630
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   511 atgacttcagattatgaattatcatttctcgccttcctcaatttggggtgcac 570
DB   631 www.wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 690
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   571 tcagagattgtgattggaacaatgttgt 601
DB   691 www.wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 721
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT      3
ID          AAF58254
AC          AAF58254 standard; DNA; 936 BP.
XX
XX          AAF58254;
XX
XX          24-APR-2001 (first entry)
XX
XX          Oligonucleotide D1875.
XX

```

KM Electron-transfer group; ETM; mismatch; genotyping;
 KW gene expression; ss.
 XX Synthetic.
 OS
 PN WO200107665-A2.
 XX
 PD 01-FEB-2001.
 XX
 XX 26-JUL-2000; 2000WO-US20476.
 PF
 XX 26-JUL-1999; 99US-0145695.
 PR 17-MAR-2000; 2000US-0190259.
 XX
 PA (CLIN-) CLINICAL MICRO SENSORS INC.
 XX
 PI Uneek RM;
 XX
 DR WPI; 2001-159728/16.
 XX
 PT Nucleic acids containing electron-transfer group, useful as labels in hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface
 XX
 PS Example 6; Page 127; 159pp; English.
 XX
 CC The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETM) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) CC and single-nucleotide polymorphisms, e.g. for genotyping, monitoring gene expression.
 CC
 XX
 SQ Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;

Query Match 15.6%; Score 94.2; DB 22; Length 936;
 Best Local Similarity 1.1%; Pred. No. 5.2e-16;
 Matches 6; Conservative 356; Mismatches 209; Indels 0; Gaps 0;
 QY 31 ttcttgatttcctccagaatacactgctcagataatgagtcacagacttcagcc 90
 DB 151 www.
 QY 91 acgaccttccaactcaagcccttcgcaaaatatttctgagaataatgaaatctta 150
 DB 211 www.
 QY 151 gggactatccagatcctgttggaataatgaaccttcttggagttatcttccttc 210
 DB 271 www.
 QY 211 acctgttaaacatatacgaagttcccttataatcttccttcagatatccattctg 270
 DB 331 www.
 QY 271 ggcctgtttgttcataatctcgagccttcctaatctgcagtgaaagaaacacaca 330
 DB 391 wgctta.
 QY 331 gaacctgtataatgaagcgaataatgaatcttctagtcgctgagagcaatagct 390
 DB 451 www.
 QY 391 ggaatcctctctacattgtttcactgatacaaaactacattgtgttattct 450
 DB 511 www.
 QY 451 caccaaataagtcagtgtaagcgttactgctgtcttcttggaatttgattacattg 510
 DB 571 www.
 QY 511 atgacttcagcattatgaattatcattctctgccttcacatttggtggcgcaac 570

DB 631 www.
 QY 571 tcagagattgtgattgtgaacaatgttgt 601
 DB 691 www.
 RESULT 4
 ID AAF58257 standard; DNA; 936 BP.
 XX
 AC AAF58257;
 XX
 DT 24-APR-2001 (first entry)
 XX
 DE Oligonucleotide D1954.
 XX
 KM Electron-transfer group; ETM; mismatch; genotyping;
 KW gene expression; ss.
 OS Synthetic.
 OS
 PN WO200107665-A2.
 XX
 PD 01-FEB-2001.
 XX
 XX 26-JUL-2000; 2000WO-US20476.
 PF
 XX 26-JUL-1999; 99US-0145695.
 PR 17-MAR-2000; 2000US-0190259.
 XX
 PA (CLIN-) CLINICAL MICRO SENSORS INC.
 XX
 PI Uneek RM;
 XX
 DR WPI; 2001-159728/16.
 XX
 PT Nucleic acids containing electron-transfer group, useful as labels in hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface
 XX
 PS Example 6; Page 127; 159pp; English.
 XX
 CC The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETM) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) CC and single-nucleotide polymorphisms, e.g. for genotyping, monitoring gene expression.
 CC
 XX
 SQ Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other;
 Query Match 15.6%; Score 94.2; DB 22; Length 936;
 Best Local Similarity 1.1%; Pred. No. 5.2e-16;
 Matches 6; Conservative 356; Mismatches 209; Indels 0; Gaps 0;
 QY 31 ttcttgatttcctccagaatacactgctcagataatgagtcacagacttcagcc 90
 DB 151 www.
 QY 91 acgaccttccaactcaagcccttcgcaaaatatttctgagaataatgaaatctta 150
 DB 211 www.
 QY 151 gggactatccagatcctgttggaataatgaaccttcttggagttatcttccttc 210
 DB 271 www.
 QY 211 acctgttaaacatatacgaagttcccttataatcttccttcagatatccattctg 270
 DB 331 www.

Oy	271	ggcctcttggcctaattccggagccttcctaattgcagltgaagaagaacaca	330
Dd	331	wgcttaaa	450
Oy	331	gaacctcgataataggcgcgaataatgatctcttagtgccctgagagcaatagct	390
Dd	451	aa	510
Oy	331	ggatcatcttcctcacatttgtttcatcctgatcaaacactaatgttggtattct	450
Dd	511	aa	570
Oy	451	caccaaatagtcaagttaaagcgttactgcctctcttcttggaatttgatatcatg	510
Dd	571	aa	630
Oy	511	atgacttcgacgataltgatattatctctctgccttcctcaatttggggtcac	570
Dd	631	aa	690
Oy	571	tcaaggatctgtgattgttgacaatgttgtt	601
Dd	691	aaa	721
 RESULT 5 AAF58259 standard; DNA; 936 BP.			
XX	AAFS8259;		
AC			
XX			
DT	24-APR-2001 (first entry)		
XX			
DE	Oligonucleotide D2004.		
KW	Electron-transfer group; ETM; mismatch; genotyping; gene expression; ss.		
OS	Synthetic.		
PN	WO200107665-A2.		
XX			
PD	01-FEB-2001.		
PF	26-JUL-2000; 2000MO-US20476.		
XX			
XX	26-JUL-1999; 99US-0145695. 17-MAR-2000; 2000US-0190259.		
PA	(CLIN-) CLINICAL MICRO SENSORS INC.		
PI	Umek RM;		
DR	WPI; 2001-159728/16.		
PT	Nucleic acids containing electron-transfer group, useful as labels in hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface		
PS	Example 6; Page 128; 159pp; English.		
CC	The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETM) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) and single-nucleotide polymorphisms, e.g. for genotyping, monitoring gene expression.		
SO	Sequence 936 BP; 6 A; 138 C; 8 G; 8 T; 776 other;		
Query Match	15.6%; Score 94.2; DB 22; Length 936;		

```

Best Local Similarity 1.1%; Pred No. 5 2e-16;
Matches 6; Conservative 356; Mismatches 209; Indels 0; Gaps 0;
OY 31 ttctctgtatttcctccagaaatcacctcgtctcagaatatgagtcacagacttcagcc 90
   :: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 151 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 210
OY 91 acgaccttctcaactcaagcccttcgaaaaattatttgcagataaataaactcta 150
   . . . : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 211 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 270
OY 151 gggactaccagatcctgtttggaataagaccttctcttggaattatcttccttc 210
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 271 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 330
OY 211 acctgttaaaacccatccaaagttcccttataattcttcagagatccattctg 270
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 331 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 390
OY 271 ggcctgttttgcattcaatctcgagccttcctaattgcagtgaaagaaacacaca 330
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 391 wgctctaa.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 450
OY 331 gaacctctgataatatgagccgaataatgactctcttagtgcctcgagagcaatagct 390
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 451 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 510
OY 391 ggaatcattccctcaacatttggttccatccagatacaaaactacattgtgtattct 450
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 511 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 570
OY 451 caaccaaatagtcagtgtaagcgtgtactcgttcctgtctcttggaatttgattacattg 510
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 571 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 630
OY 511 atgacttcagcatattgaattatcaattctctcgccttcctaatttgggtgcac 570
   . . . : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 631 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 690
OY 571 tcagagattgtgatgtcgaaacaatgttgt 601
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 691 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 721

RESULT 6
AAFS8262
AAFS8262 standard; DNA; 936 BP.
AC AAF58262;
XX
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D2007.
XX
KW Electron-transfer group; ETM; mismatch; genotyping;
   gene expression; ss.
XX
XX Synthetic.
XX
XX WO200107665-A2.
XX
XX PN 01-FEB-2001.
XX
XX PD 26-JUL-2000; 2000WO-US20476.
XX
XX PE 26-JUL-1999; 99US-0145695.
XX
XX PR 17-MAR-2000; 2000US-0190259.
XX
XX PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
XX PT Umek RM;
XX
XX WPI; 2001-159728/16.
XX
XX DR

```

QY 511 atgacttcagcattattgaatttctctcgtccttctcaatttgggtgccac 570

QY 511 atgacttcagcattattgaatttctctcgtccttctcaatttgggtgccac 570

Db 631 wwwwww
Qy 571 tcagagatlgatltgtaacaatgtgtt 601
Db 691 wwwwww
RESULT 8
AAF58252/c
ID AAF58252 standard; DNA; 936 BP.
XX
AC AAF58252;
XX
XX 24-APR-2001 (first entry)
XX
DE Oligonucleotide D1835.
XX
XX Electron-transfer group; ETM; mismatch; genotyping;
KM gene expression; ss.
XX
XX Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
XX 26-JUL-2000; 2000MO-US20476.
XX
PF 26-JUL-1999; 99US-0145695.
XX
PR 17-MAR-2000; 2000US-0190259.
XX
XX (CLIN-) CLINICAL MICRO SENSORS INC.
XX
XX Umek RM;
XX
XX WPI; 2001-159728/16.
XX
XX Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface -
XX
XX Example 6; Page 127; 159pp; English.
XX
XX The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
SQ Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;
Query Match 15.5%; Score 93.4; DB 22; Length 936;
Best Local Similarity 0.7%; Pred. No. 8.5e-16;
Matches 4; Conservative 356; Mismatches 209; Indels 0; Gaps 0;
Qy 31 ttctgtattcttcgcagaatcactgcttcagaatagtcacagaacttcagcc 90
Db 675 wwwwww
Qy 91 acgaactttcaactcaagcccttcgcaaaatattgtctagaanaaataccta 150
Db 615 wwwwww
Qy 151 gggacttcagatccgttggatgaatagactttcttggagtatacttccttcc 210
Db 555 wwwwww
Qy 211 acctgttaaacatatccaaagttcccttatattcttcagatatcatctcgg 270
Db 495 wwwwww

Qy 271 ggcctgttttctaatcctcgagccttcctcaattgcagtgaaagaaacacca 330
Db 435 wwwwww
Qy 331 gaaactcgtataatctgagccgagataatgaatcttccttagtgccctgagagcaatgct 390
Db 375 wwwwww
Qy 391 ggaatccttcctcacaattggttccctcctagatacaaaacatttgygtattct 450
Db 315 wwwwww
Qy 451 caccacaatagtcagtgtaaggcgttactgctcgttcttgaggaaattgatcata 510
Db 255 wwwwww
Qy 511 atgaacttcagcatattgaatattcaattctcctccttcatttgagggtccac 570
Db 195 wwwwww
Qy 571 tcagagatlgatltgtaacaatgtgtt 601
Db 135 wwwwww
RESULT 9
AAF58254/c
ID AAF58254 standard; DNA; 936 BP.
XX
AC AAF58254;
XX
XX 24-APR-2001 (first entry)
XX
DE Oligonucleotide D1875.
XX
XX Electron-transfer group; ETM; mismatch; genotyping;
KM gene expression; ss.
XX
XX Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
XX 26-JUL-2000; 2000MO-US20476.
XX
PF 26-JUL-1999; 99US-0145695.
XX
PR 17-MAR-2000; 2000US-0190259.
XX
XX (CLIN-) CLINICAL MICRO SENSORS INC.
XX
XX Umek RM;
XX
XX WPI; 2001-159728/16.
XX
XX Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface -
XX
XX Example 6; Page 127; 159pp; English.
XX
XX The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
SQ Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;
Query Match 15.5%; Score 93.4; DB 22; Length 936;
Best Local Similarity 0.7%; Pred. No. 8.5e-16;

Matches 4; Conservative 358; Mismatches 209; Indels 0; Gaps 0;

QY 31 ttctgtattctccccaagaatacactgcttcagaataatggtccagaacttcgcc 90
Db 675 www 616
QY 91 acgaccttcaactcaaaagcccttcgcaaaattattgtagaaaaatgaatctta 150
Db 615 www 556
QY 151 gggactccagatccgtttggaataatgaacttcttcttgagttatctcttc 210
Db 555 www 496
QY 211 acctgttaaacacataccaagttcccttattcttcttcagatatccattctg 270
Db 495 www 436
QY 271 ggcctgtttgtcattatctcgagcccttctaattgcagtgaagaagaacaca 330
Db 435 www 376
QY 331 gaactctgaataatagagccgaataatgaatcttcttgaagccctgagagcaatgct 390
Db 375 www 316
QY 391 ggaactatccctccacattgtttcatcctagatcaaaactcattgtgtattct 450
Db 315 www 256
QY 451 caccaaatagtcagtgtaagcgtctactgctctgtcttggaatttgattcatg 510
Db 255 www 196
QY 511 atgacttcagcattatgaattatcattctctgccttcctaatttggggtgcac 570
Db 195 www 136
QY 571 tcagagatgtgattgtgaacaatgttgtt 601
Db 135 ww 105

RESULT 10
AAF58257/C
ID AAF58257 standard; DNA; 936 BP.
XX
AC AAF58257;
XX
24-APR-2001 (first entry)
XX
DE Oligonucleotide D1954.
XX
KM Electron-transfer group; ETM; mismatch; genotyping;
XX
OS Synthetic.
XX
WO200107665-A2.
XX
PN
XX
01-FEB-2001.
XX
PD
XX
26-JUL-2000; 2000MO-US20476.
XX
PF
XX
26-JUL-1999; 990US-0145695.
PR
XX
17-MAR-2000; 2000US-0190259.
XX
PA
XX
(CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI
XX
Imek RM;
XX
WPI; 2001-159728/16.
XX

PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface
XX
XX
PS Example 6; Page 127; 159pp; English.
XX
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
SQ Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other;

Query Match 15.5%; Score 93.4; DB 22; Length 936;
Best Local Similarity 0.7%; Pred. No. 8.5e-16;
Matches 4; Conservative 358; Mismatches 209; Indels 0; Gaps 0;

QY 31 ttctgtattctccccaagaatacactgcttcagaataatggtccagaacttcgcc 90
Db 675 www 616
QY 91 acgaccttcaactcaaaagcccttcgcaaaattattgtagaaaaatgaatctta 150
Db 615 www 556
QY 151 gggactccagatccgtttggaataatgaacttcttcttgagttatctcttc 210
Db 555 www 496
QY 211 acctgttaaacacataccaagttcccttattcttcttcagatatccattctg 270
Db 495 www 436
QY 271 ggcctgtttgtcattatctcgagcccttctaattgcagtgaagaagaacaca 330
Db 435 www 376
QY 331 gaactctgaataatagagccgaataatgaatcttcttgaagccctgagagcaatgct 390
Db 375 www 316
QY 391 ggaactatccctccacattgtttcatcctagatcaaaactcattgtgtattct 450
Db 315 www 256
QY 451 caccaaatagtcagtgtaagcgtctactgctctgtcttggaatttgattcatg 510
Db 255 www 196
QY 511 atgacttcagcattatgaattatcattctctgccttcctaatttggggtgcac 570
Db 195 www 136
QY 571 tcagagatgtgattgtgaacaatgttgtt 601
Db 135 ww 105

RESULT 11
AAF58259/C
ID AAF58259 standard; DNA; 936 BP.
XX
AC AAF58259;
XX
24-APR-2001 (first entry)
XX
DE Oligonucleotide D2004.
XX
KM Electron-transfer group; ETM; mismatch; genotyping;
XX
gene expression; ss.
XX

```
XX Synthetic.
OS
XX WO200107665-A2.
PN
XX 01-FEB-2001.
PD
XX 26-JUL-2000; 2000WO-US20476.
PF
XX 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX
XX (CLIN-) CLINICAL MICRO SENSORS INC.
XX
XX Umek RM:
XX
XX WPI; 2001-159728/16.
XX
XX Nucleic acids containing electron-transfer group, useful as labels in
XX hybridization assays, e.g. for genotyping, allowing repeat analyses on
XX a single surface
XX
XX Example 6; Page 128; 159pp; English.
XX
XX The present invention relates to a composition comprising two nucleic
XX acids each containing an electron-transfer group (ETM) having
XX different redox potentials. The invention is used for electronic
XX detection of nucleic acids, especially of substitutions (mismatches)
XX and single-nucleotide polymorphisms, e.g. for genotyping,
XX monitoring gene expression.
XX
XX Sequence 936 BP; 6 A; 138 C; 8 G; 8 T; 776 other:
SQ

Query Match          15.5%; Score 93.4; DB 22; Length 936;
Best Local Similarity 0.7%; Pred. NO. 8.5e-16;
Matches 4; Conservative 350; Mismatches 209; Indels 0; Gaps 0;

OY 31 ttctgtatttcctccagaatacactgcttcagaaatagatcagacacttcagcc 90
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 675 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 616

OY 91 acgaccccttccaactcaagccctcgcaaaaattattgctagaataatgaaatccta 150
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 615 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 556

OY 151 gggactatccagatccgtttggaattagaccttcttttggaattacttcctttc 210
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 555 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 496

OY 211 acctgttaaacacataccaaggttccttataattcttccaagataccactctg 270
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 495 wgwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 436

OY 271 ggcctgtttgttcataattcgtgagccttctaattgcagtgaataaaacacaca 330
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 435 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 376

OY 331 gaactgtataatagagccgaataatgaatcttcttagtgcccgagagcaatagct 390
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 375 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 316

OY 391 ggaatcattcctccacattgtttcatcctagatcaaaactacattgtgtattct 450
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 315 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 256

OY 451 caccacaaatgctcagtgtaagctgtactgtcctgttcttggaatttgatatactg 510
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 255 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 196

OY 511 atgacttcagcattatgataatattcattctcctgccttccaatttgggggtgcac 570
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 195 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 136
```

```
OY 571 tcagagagattgtatgtgacaaatgttgt 601
   : : : : : : : : : : : : : : : : : :
Db 135 wwwwwwwwwwwwwwwwwwwwwwwwwwww 105

RESULT 12
AAF58262/c
ID AAF58262 standard; DNA; 936 BP.
XX
XX AAF58262;
AC
XX 24-APR-2001 (first entry)
DT
XX Oligonucleotide D2007.
DE
XX
XX Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX
XX Synthetic.
OS
XX WO200107665-A2.
PN
XX 01-FEB-2001.
PD
XX 26-JUL-2000; 2000WO-US20476.
PF
XX 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX
XX (CLIN-) CLINICAL MICRO SENSORS INC.
XX
XX Umek RM:
XX
XX WPI; 2001-159728/16.
XX
XX Nucleic acids containing electron-transfer group, useful as labels in
XX hybridization assays, e.g. for genotyping, allowing repeat analyses on
XX a single surface
XX
XX Example 6; Page 128; 159pp; English.
XX
XX The present invention relates to a composition comprising two nucleic
XX acids each containing an electron-transfer group (ETM) having
XX different redox potentials. The invention is used for electronic
XX detection of nucleic acids, especially of substitutions (mismatches)
XX and single-nucleotide polymorphisms, e.g. for genotyping,
XX monitoring gene expression.
XX
XX Sequence 936 BP; 5 A; 139 C; 10 G; 6 T; 776 other:
SQ

Query Match          15.5%; Score 93.4; DB 22; Length 936;
Best Local Similarity 0.7%; Pred. NO. 8.5e-16;
Matches 4; Conservative 358; Mismatches 209; Indels 0; Gaps 0;

OY 31 ttctgtatttcctccagaatacactgcttcagaaatagatcagacacttcagcc 90
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 675 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 616

OY 91 acgaccccttccaactcaagccctcgcaaaaattattgctagaataatgaaatccta 150
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 615 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 556

OY 151 gggactatccagatccgtttggaattagaccttcttttggaattacttcctttc 210
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 555 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 496

OY 211 acctgttaaacacataccaaggttccttataattcttccaagataccactctg 270
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 495 wgwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 436

OY 271 ggcctgtttgttcataattcgtgagccttctaattgcagtgaataaaacacaca 330
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 195 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 136
```



```

PR 04-SEP-1998: 98US-00992229.
PR 23-OCT-1998: 98US-0105368.
PR 08-JAN-1999: 99US-0115234.
PR 12-FEB-1999: 99US-0119931.
PR 18-FEB-1999: 99US-0120575.
PR 30-APR-1999: 99US-0132020.
PR 11-AUG-1999: 99US-0096622.
XX
XX (GENM ) GENETICS INST INC.
XX
PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M, Agostino MJ, Steininger RJ, Spaulding V;
PI Wong GG, Clark HF, Fechtel K;
DR WPI: 2000-205979/18.
DR P-PSDB: AAY94973.
XX
XX New polynucleotides encoding secreted proteins, which may have e.g.
XX nutritional, chemokine, immune stimulating or suppressing,
XX hematopoiesis regulating, tissue growth, activin/inhibin
XX antiinflammatory or tumor inhibition activity -
XX
PS Claim 160: Page 606-607; 641pp; English.
XX
XX AA16618 to AA16697 encode the human secreted proteins given in
CC AAY94898 to AAY94980, isolated from human adult brain, adult thyroid,
CC adult retina, foetal carcinoma, adult blood, adult neural, foetal kidney,
CC adult placenta, adult testis, whole embryo, adult cartilage, kidney,
CC foetal brain, adult thymus, foetal placenta, adult uterus, adult tumour,
CC and adult bladder, cDNA libraries. The polynucleotides and proteins are
CC predicted to have biological activities which would make them suitable
CC for treating, preventing or ameliorating medical conditions in humans
CC and animals. The polynucleotides can be used as markers for tissues in
CC which the protein is preferentially expressed, as molecular weight
CC markers on Southern gels, and as chromosome markers or tags to identify
CC chromosomes or to map gene positions. The proteins can be used in the
CC treatment of immune deficiencies and disorders, such as severe combined
CC immunodeficiency (SCID), as well as viral, bacterial, fungal and other
CC infections. These infections include human immunodeficiency virus (HIV),
CC hepatitis, herpesviruses, mycobacteria, leishmania spp., malaria and
CC candidiasis. The proteins can be used to treat autoimmune disorders such
CC as connective tissue disease, multiple sclerosis, systemic lupus
CC erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation,
CC Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent
CC diabetes mellitus, myasthenia gravis, graft-versus-host-disease and
CC autoimmune inflammatory eye disease. The proteins can also be used to
CC treat allergic conditions, such as asthma. AA16698 to AA16774 represent
CC probes for the human secreted proteins from the present invention.
SQ
SQ Sequence 1330 BP; 364 A; 285 C; 309 G; 372 T; 0 other;

```

```

Query Match 14.3%; Score 86.4; DB 21: Length 1330;
Best Local Similarity 51.1%; Pred. No. 7.2e-14;
Matches 232; Conservative 0; Mismatches 216; Indels 6; Gaps 1;

```

```

OY 119 aaaaattattgttagaaaatgaanaatcttagagactaccgacatcttggaa178
DB 485 agaagttcttgaaggagagaccacagctccttggggttgagcagattctgacgcctga 544
OY 179 tgaaccttcttcttgagatattctcttccacttgtaaaacacatcaaggttc 238
DB 545 tggagcttgatcgatggaataacaatgatgtgtatgcatcattactatcagaagtaac 604
OY 239 ccttataattcttcaagatcattctgggctctgtttgttcataattctggag 238
DB 605 ctatttcgtatatacgggttacacaatttgggggtcagaatgaattattatcagat 654
OY 299 ccttctaattgagagaagaagaacacacgaactcgcataaattgagccgaataa 358
DB 665 ccttgccaattgagcagaagaataagaactcaaaaagcctggtccgaggtgagctaggaa 724
OY 359 tgaatcttcttagtccttgaagagcaatagctggaatcattctctcctcattgtgttca 418

```

```

DB 725 tgatatcaccagctcgtactgctgcacacagagatcttaacacattagcttg 784
OY 419 tctcagat-----cacaactacattgtgttattctcccaaaatagcagtgtaagg 472
DB 785 cgtttattcaccacacccttaccgttaacacataggaactcaaaatgttcatg 844
OY 473 ctgtactgcctgtcttcttggaatttgattacattgaacttcagcattatgaat 532
DB 845 ggaactatgcatacttaatgagtgctgtagtgatgagcgtgcctccttaagtgtcgtgaat 904
OY 533 tatctattctcgtccttcttcaatttggggtg 566
DB 905 tctgactgtgtgtccctctcgtccttggatg 938

```

RESULT 15

```

AAK15104
ID AAK15104 standard; cDNA; 1669 BP.
XX
AC AAK15104;
XX
XX 19-APR-1999 (first entry)
XX
XX
XX High affinity immunoglobulin E receptor-like protein (IGERB) cDNA.
XX
XX High affinity immunoglobulin E receptor-like protein: IGERB;
KW Incyte clone 927955; inflammatory response; AIDS; Addison's disease;
KW atherosclerosis; bronchitis; ulcerative colitis; diabetes mellitus;
KW emphysema; gout; Graves's disease; osteoporosis; rheumatoid arthritis;
KW Sjogren's syndrome; cancer; ss.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX CDS 159..878
XX FT /*tag= a
XX FT /product= IGERB
XX
XX US5871930-A.
XX
XX 16-FEB-1999.
XX
XX 21-AUG-1997; 97US-0916902.
XX
XX 21-AUG-1997; 97US-0916902.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Bandman O, Corley NC, Lal P;
XX
XX WPI: 1999-166628/14.
XX
XX P-PSDB: AAM96745.
XX
XX
XX New polynucleotide coding for high affinity immunoglobulin E
XX receptor-like protein - and its complement, useful for diagnosis,
XX prevention and antisense therapy of inflammatory responses
XX
XX Claim 4: Fig 1A-D: 30pp; English.
XX
XX
XX The present sequence encodes a high affinity immunoglobulin E
XX receptor-like protein (IGERB). The present sequence was first
XX identified in Incyte clone 927955 for the brain cDNA library
XX BRAIN0104. Polynucleotides complementary to the IGERB cDNA can
XX be used as probes to IGERB gene expression. IGERB proteins and
XX nucleotides may be used for the diagnosis of inflammatory
XX responses associated with expression of IGERB, e.g AIDS,
XX Addison's disease, atherosclerosis, bronchitis, ulcerative colitis,
XX diabetes mellitus, emphysema, gout, Graves's disease, osteoporosis,
XX rheumatoid arthritis, Sjogren's syndrome and complications of cancer.
XX The sequences may also be useful in assays that detect activation or
XX induction of various cancers.
XX
XX

```

SQ Sequence 1669 BP; 481 A; 350 C; 334 G; 501 T; 3 other;

Query Match 14.3%; Score 86.4; DB 20; Length 1669;
Best Local Similarity 51.1%; Pred. No. 7.8e-14;
Matches 232; Conservative 0; Mismatches 216; Indels 6; Gaps 1;

```
QY 119 aaaaattattgtagaanaatgaanaatcttaggactaccagatcctgttgaatta 178
Db 325 agaagttctgaaggaggaacccaagtccttgggttgcagatctgcacgcctga 384
QY 179 tgacctttcttttggagttatcttcctttcaccttgtaaaacacatccaagttc 238
Db 385 tgaagccttagcatggygaataacaatgatgtatgcatctaatctaatgagaatacc 444
QY 239 cctttatattcttcaaggatcaccatctcgggctcgtttgttcataattctggag 298
Db 445 ctatttcggtatatacgggtacacaatttgggggtcagtaagttatatttcaggat 504
QY 299 ccttcctaattgcagtgaaagaaacacagaaactcgtataatttgagcgaataa 358
Db 505 ccttgcctaattgcagcaggaattagaactacaaagcctgtgccgagtgagctagaa 564
QY 359 tgaatcttcttagtgccctgaagagaatagcttgaatcaltctcctcacatttggttca 418
Db 565 tgaatatcaccaagctctgtactgctgcatacaggatcttaatcaacacatttagcttg 624
QY 419 tccctagat-----caaaactacatttgggtatctcctcaaaaatagtcagtgaa 472
Db 625 cgtttatcattccatcacacccttactgttaactatgtgcaactcaataatgtcatg 684
QY 473 ctgttactgcctgtcttcttgggaatttgaattgaatgaatgaacttcaagcattatgaat 532
Db 685 ggaatatgtccatcattaaatgggtcgtgatgcatgtgtcctccttaagtgctggaat 744
QY 533 tattcaattctctgccttctcctaatttgggggtg 566
Db 745 tctgcaatgctgtctcctctcgtccttggatcg 778
```

Search completed: October 7, 2001, 04:07:30
Job time: 2810 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 7, 2001, 03:19:45 ; Search time 37.7 Seconds
(without alignments)
3027.974 Million cell updates/sec

Title: US-09-735-712-1

Perfect score: 603

Sequence: 1 atgattcaagcaccgcaca.....attgtgaacatgtttga 603

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues 649198

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/prodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/prodata/2/lna/6A.COMB.seq:*
4: /cgn2_6/prodata/2/lna/6B.COMB.seq:*
5: /cgn2_6/prodata/2/lna/PCrUS.COMB.seq:*
6: /cgn2_6/prodata/2/lna/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86.4	14.3	1669	2	US-08-916-902A-2
2	86.4	14.3	1669	2	US-09-213-389-2
3	63.2	10.5	1661	1	US-08-318-492-3
4	63.2	10.5	1661	1	US-08-707-340-3
5	63.2	10.5	1661	2	US-08-994-578-3
6	49.2	8.2	2545	1	US-07-869-933-22
7	49.2	8.2	2545	4	US-09-103-663-22
8	42.6	7.1	7218	1	US-08-732-463-14
9	34.6	5.7	246240	2	US-08-724-394A-20
10	34.6	5.7	246240	2	US-08-724-394A-21
11	34.6	5.7	246240	2	US-08-724-394A-22
12	33.6	5.6	1711	1	US-08-568-147B-1
13	33.4	5.5	6669	3	US-09-212-971-5
14	33.4	5.5	6669	4	US-08-800-929A-5
15	32.2	5.3	3947	4	US-08-975-762-47
16	32.2	5.3	8920	4	US-08-446-855A-1
17	32	5.3	8920	4	US-09-150-741-1
18	32	5.3	9793	1	US-08-470-202-56
19	32	5.3	9793	1	US-08-471-770-56
20	32	5.3	9793	2	US-08-468-059-56
21	31.4	5.2	3393	4	US-09-104-324B-1
22	31.4	5.2	3393	4	US-09-162-713-1
23	31.4	5.2	3580	4	US-09-081-345-1
24	31.4	5.2	4291	2	US-08-417-210A-81
25	31.4	5.2	4517	4	US-09-140-804-9
26	31	5.1	1502	4	US-08-793-044-4
27	31	5.1	2839	4	US-09-061-702-1

28	31	5.1	2945	3	US-09-058-489-85	Sequence 85, Appl
29	31	5.1	9439	3	US-09-058-489-89	Sequence 89, Appl
30	30.8	5.1	669	5	PCT-US95-13335-2	Sequence 2, Appl1
31	30.8	5.1	2662	3	US-08-750-357-8	Sequence 8, Appl1
32	30.8	5.1	3466	2	US-08-468-036-38	Sequence 38, Appl1
33	30.6	5.1	4526	1	US-07-855-412B-4	Sequence 4, Appl1
34	30.6	5.1	4526	2	US-08-308-887A-4	Sequence 4, Appl1
35	30.6	5.1	4526	3	US-08-881-094-4	Sequence 4, Appl1
36	30.4	5.0	2189	3	US-08-846-020A-1	Sequence 1, Appl1
37	30.4	5.0	3427	3	US-09-009-913-2	Sequence 2, Appl1
38	30.4	5.0	5510	3	US-09-009-913-3	Sequence 3, Appl1
39	30.4	5.0	5667	3	US-09-009-913-4	Sequence 4, Appl1
40	30.4	5.0	22846	2	US-08-469-461-3	Sequence 3, Appl1
41	30.4	5.0	22846	3	US-07-890-609-3	Sequence 3, Appl1
42	30.2	5.0	315	2	US-08-743-200-1	Sequence 1, Appl1
43	30.2	5.0	376	2	US-08-623-906A-18	Sequence 18, Appl
44	30.2	5.0	1956	1	US-08-318-831-7	Sequence 7, Appl1
45	30	5.0				

ALIGNMENTS

RESULT 1
US-08-916-902A-2
; Sequence 2, Application US/08916902A
; Patent No. 5871930
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
; TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/916,902A
; FILING DATE: Herewith
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT04
; CLONE: 927955
; US-08-916-902A-2

```

TELEFAX: 415-845-4166
:
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
:   LENGTH: 1669 base pairs
:   TYPE: nucleic acid
:   STRANDEDNESS: single
:   TOPOLOGY: linear
: IMMEDIATE SOURCE:
:   LIBRARY: BRAINOT04
:   CLONE: 927955
:
: US-09-213-389-2
:
Query Match      14.3%  Score 86.4:  DB 2:  Length 1669;
Best Local Similarity 51.1%:  Pred. No. 5.9e-15;
Matches 232:  Conservative 0:  Mismatches 216;  Indels 6;  Gaps 1;

QY  119  aaaaattatttgcctagaagaaaatgaacatctagaggactatccagatccctgtttgaa178
Db   325  AGAAGTTCTTGAAGGGAGAACCCAAAGTCCTTGGGGTGTGCAGATTTGACTGCGCCGGA 384
QY  179  tgacctttctcttggaggtatctctcttccattcacttggttaaacatatccaggtttc 238
Db   385  TGAGCCTTTCAGCATGGGAATTAACAATGATGTGTATGGCATCTAATTAATGGAAGTAAC 444
QY  239  ccttataattcttccaagatatccattctggggctcgtttgttttcataattcggag 298
Db   445  CTATTCGCGTGAATATCGGGTATACAAATTTGGGGGTCAGTAAATGTTATTTTCAGAT 504
QY  299  ccttctaatattgcagtgtaaaagaacacagaacatctgataattatggccgaataa 358
Db   505  CTTGTCTCAATTTGACACAGGAATTAACAATAACAAAAGGCTGTGGTCCGAGGTAGTGA 564
QY  359  tgaatcctcttagtgccctcgagagagcaatagctggaatcatctccctacatttg1tca 418
Db   565  TGAATATACACAGCTCTGTACTGTGCTGATCAGGATCTTAATCAACATTTAAGCTTGG 624
QY  419  tcttagat-----caaaactacattgtgttattctcaccaaaatagtcagtgtaagg 472
Db   625  CGTTTATTTCATTCATTCATCACCTTACTGTACTACTAATGAGCAACATAATTAATGTG 684
QY  473  ctgttaacgtctcgttcttcttggaatttgatcaattgatgaacttgcacttat1tgat 532
Db   685  GGACTATGTCCATCTTAATAGGGTCTGTGATGGCATGTGTCCTCTTAAGTGTGCGAAT 744
QY  533  taticattctctgccttcttccaaatl1tgg99gtg 566
Db   745  TCTGCATTCGTGTGTCCCTCTCTGCTGCTTGGATG 778

RESULT 3
US-08-318-492-3
: Sequence 3, Application US/08318492
: Patent No. 5552312
: GENERAL INFORMATION:
:   APPLICANT: Lim, Bing
:   APPLICANT: Adra, Chaker N.
:   APPLICANT: Lellias, Jean-Michel
:   TITLE OF INVENTION: RECOMBINANT HTM4 GENE, PROTEIN AND ASSAYS
:   NUMBER OF SEQUENCES: 4
:   CORRESPONDENCE ADDRESS:
:     ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
:     STREET: Two Millitia Drive
:     CITY: Lexington
:     STATE: Massachusetts
:     COUNTRY: U.S.A.
:     ZIP: 02173
:   COMPUTER READABLE FORM:
:     MEDIUM TYPE: Floppy disk
:     COMPUTER: IBM PC compatible
:     OPERATING SYSTEM: PC-DOS/MS-DOS

```



```

QY 6 tccaagacgcacagatccggtgttctgtatctccacagaatcactgcttcaaga 65
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 875 TTAAGATTTCATACACAGCAGTAGTTTACACCATTTCCAGATTAATGATAAGA 934
QY 66 atagatgccacagacttcagccagactttcaactgaagccctcgcaaaatt 125
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 935 TTCAAAATATTATTAAACGGTTTACGTTGAATGATGCCATGAGTCGGCTACTATACT 994
QY 126 attgctagaaaaaataaactataggaactatccagatcctgttgaattagacct 185
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 995 ATTTCCTCTGTTGCATACGCTCACAGATAATTATCCAGCTTGCTGCAGTCAGTCAG 1054
QY 186 tctcttggaatctctctctcttccactgtttaaacacatccaggttcccttat 245
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1055 GAGACTTGCGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1114
QY 246 attcttccagatatacatctcgtggtctgttcttcttcaatcttgagacttcc 305
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1115 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1174
QY 306 aattgcagtgaaaagaaacacagaaactctgataataltgagccgaataatc 365
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1175 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1234
QY 366 tcttagtgcctgagagcaatagctgaatcatctcctccacatttggttcaactaga 425
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1235 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1294
QY 426 tcaaaactacattgttgtaattctccacaaatagtcagtgtaagctgttactgcct 485
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1295 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1354
QY 486 gtcttggaatttgatcatatgactgactcagcatatgaattatcattctct 545
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1355 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1414
QY 546 gccttctcaattt 560
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1415 YYYYYYYYYYYYYYYY 1429

```

RESULT 9

```

US-08-724-394A-20
; Sequence 20, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Krommal, Gregory S.
; APPLICANT: Lauder, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchinashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
; TITLE OF INVENTION: Sequences and Antibodies Thereeto
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:

```

```

NAME: Fitts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 017957-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 246240 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1..246240
OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-20

```

```

Query Match 5.7%; Score 34.6; DB 2; Length 246240;
Best Local Similarity 55.4%; Pred. No. 6.8;
Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

```

```

QY 449 ctccacaaatagtcagtgtaaggctgttactgtcctgttcttggaatttgatcat 508
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36590 CTCCTCAGATCTCGTATTATTCATTCATTTATGCTGTGTAGTAAATACCAACTG 36749
QY 509 tgatgaattcagcatatgaattatcatctctcgccttctcaatttggggcgc 568
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36750 TGATGGCTTGAAACACCAAAATATTATTATTACAAATCTAAAGCTCAGAACTCC 36809
QY 569 a 569
Db 36810 A 36810

```

RESULT 10

```

US-08-724-394A-21
; Sequence 21, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Krommal, Gregory S.
; APPLICANT: Lauder, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchinashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
; TITLE OF INVENTION: Sequences and Antibodies Thereeto
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:

```

; INFORMATION FOR SEQ ID NO: 22:

LENGTH: 1711 base pairs
TYPE: nucleic acid

TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 22
SEQUENCE CHARACTERISTICS:

Db 2886 ATGTGAGATTATTCACCTATCATCCTGTAACATTTTGTCTTCTAA 2827

Job time: 2878 sec

QY 248 ttcttcagga 258

Db 2826 ATGTTCTGAA 2816

RESULT 15

US-08-975-762-47

; Sequence 47, Application US/08975762

; Patent No. 6207169

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Lodes, Michael J.

; APPLICANT: Houghton, Raymond

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND

; NUMBER OF SEQUENCES: 73

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/975,762

; FILING DATE: 21-MAR-1997

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Maki, David J.

; REGISTRATION NUMBER: 31,392

; REFERENCE/DOCKET NUMBER: 210121.439

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 206-622-4900

; TELEFAX: 206-682-6031

; INFORMATION FOR SEQ ID NO: 47:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3947 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; US-08-975-762-47

Query Match

Best Local Similarity 5.3%; Score 32.2; DB 4; Length 3947;

Matches 103; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 365 ttcttagtgcctgaggaagcaatagcgtgcaatcctccacacattggttcacccag 424

Db 3375 TTTTGAATTTGATATGATATGTTGTGTAGTATTCACAGGCTTTCTGTTTG 3434

QY 425 atcaaaactacatttggtgattcctcacaagaatagtcagtgtaagcgtctactgtcc 484

Db 3435 ATGTTCTCTAGTTGGTGGCTGCTGTTGTAATCTATAGSTTTTATTTACACTTAAC 3494

QY 485 tgttcttggaatttgattacattgacattcagcattatgcaattatcattctc 544

Db 3495 TAAATTTGAGAGTTGTACAGCAATTAATTTCTTAATTAATTAATTTGACTTTTAAAGCCTCT 3554

QY 545 tgccttctcaatttgagggtgcacacacagagattgcat 585

Db 3555 ACATATTCATTTCTTTTGTGAGGCTCTGATGACATGAT 3595

Thu Oct 11 10:24:04 2001

us-09-735-712-1.rml

Page 10

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2001, 04:39:26 ; Search time 40.14 Seconds
(Without alignments)
225.037 Million cell updates/sec

Title: US-09-735-712-8

Perfect score: 771

Sequence: 1 MDSYHSPVFLVPPPEITA.....SLPFSILCHSEDCDECC 149

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database : A.Geneseq_0601.*
- 1: /SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT.*
 - 2: /SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT.*
 - 3: /SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT.*
 - 4: /SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT.*
 - 5: /SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT.*
 - 6: /SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT.*
 - 7: /SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT.*
 - 8: /SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT.*
 - 9: /SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT.*
 - 10: /SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT.*
 - 11: /SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT.*
 - 12: /SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT.*
 - 13: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT.*
 - 14: /SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT.*
 - 15: /SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT.*
 - 16: /SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT.*
 - 17: /SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT.*
 - 18: /SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT.*
 - 19: /SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT.*
 - 20: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT.*
 - 21: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT.*
 - 22: /SIDSI/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	294	38.1	77	AA03074	Human secreted pro
2	161.5	20.9	167	AA53632	A bone marrow secr
3	150.5	19.5	214	AA06503	Htm protein. Hom
4	150.5	19.5	214	AAW41056	Htm protein. Hom
5	139	18.0	225	AAV9449	Human inflammation
6	139	18.0	242	AAV73495	Human secreted pro
7	139	18.0	248	AAV15225	Human receptor pro
8	139	18.0	248	AAV91531	Human secreted pro
9	139	18.0	248	AAV70489	Human hHA1ERBs-iso
10	139	18.0	250	AAV48505	Human breast tumo
11	139	18.0	257	AA58419	Lung cancer associ

12	139	18.0	273	AAV91680	Human secreted pro
13	136	17.6	248	AAV36046	Extended human sec
14	134	17.4	239	AAV6745	High affinity immu
15	134	17.4	239	AAV50174	Human high affinity
16	134	17.4	239	AAV4973	Human secreted pro
17	131.5	17.1	249	AAV4584	Human HA1ERBs iso
18	131.5	17.1	249	AAV4583	Human HA1ERBs iso
19	128	16.6	267	AAV15224	Human receptor pro
20	128	16.6	269	AAV91352	Human secreted pro
21	125	16.2	307	AAV91421	Human secreted pro
22	123.5	16.0	243	AAV05026	Beta subunit of ra
23	123.5	16.0	243	AAV42341	Subunit of the hum
24	123.5	16.0	246	AAV42341	Beta subunit of h1
25	123.5	16.0	246	AAV42337	Human FCER1 beta.
26	114.5	14.9	250	AAV61619	Clone HPER86 of T
27	114.5	14.9	250	AAV70432	Human cell surface
28	114.5	14.9	302	AAV54312	Human pancreatic c
29	113.5	14.7	297	AAV0445	Human CD20.4 anti
30	113.5	14.7	297	AAV6192	Human CD20.4 anti
31	110.5	14.3	297	AAV91356	CD20.4 antigen.
32	110.5	14.3	297	AAV20808	Human CD20 antigen
33	110.5	14.3	297	AAV96131	Human cell surface
34	109.5	14.2	297	AAV91436	Human CD20 antigen
35	109	14.1	103	AAV12429	Human 5' EST secre
36	105.5	13.7	237	AAV4585	Human HA1ERBs iso
37	105.5	13.7	237	AAV70490	Human HA1ERBs prot
38	105.5	13.7	240	AAV69998	Human receptor-ass
39	99.5	12.9	195	AAV74448	Human variant Fcep
40	99.5	12.9	244	AAV29149	Human high affinity
41	99.5	12.9	244	AAV75918	Human beta subunit
42	99.5	12.9	244	AAV74447	Human wild-type Fc
43	96.5	12.5	204	AAV72900	Human IGERB SEQ ID
44	96.5	12.5	204	AAV91386	Human secreted pro
45	96	12.5	93	AAV65059	Human 5' EST relat

ALIGNMENTS

RESULT 1

AA03074 standard; Protein: 77 AA.

AA03074:

06-OCT-2000 (first entry)

Human secreted protein, SEQ ID NO: 7155.

Human: 5' EST: expressed sequence tag; secreted protein; cDNA isolation;

gene therapy; chromosome mapping.

Homo sapiens.

EP1033401-A2.

06-SEP-2000.

21-FEB-2000; 2000EP-0200610.

26-FEB-1999; 99US-0122487.

(GEST) GENSET.

Dumas Malne Edwards J, Duclert A, Giordano J;

WPI: 2000-500381/45.

N-PDB: AAC03080.

New nucleic acid that is a 5' expressed sequence tag (5' EST) for

obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for

diagnostic, forensic, gene therapy and chromosome mapping procedures -

PS Claim 13; SEQ ID 7155; 71np + CD-ROM; English.

Quency Match	38.1%	Score 294	DB 21	Length 77
Best Local Similarity	83.6%	Pred. No. 2.3e-28		
Matches	56	Conservative	3	Mismatches 8
				Indels 0
				Gaps 0

QY	93	VLFINSG	99
QY	33	FSTOSLOKIFARKMILSTIOLFGIMFESGVFLPFTTLTKPYRPRPFILSGPFWGS	92
Db	11	fqlkapockkrfarkmllgtlqxlmgmfesgvflfklxkxprlprffllslxypfwgs	70
QY	71	vlfninsg	77

RESULT	2
AAV53632	
ID	AAV53632 standard; Protein; 167 AA

DT	22-FEB-2000	(first entry)
XX		
DE	A bone marrow secreted protein designated BMS208	

KM Bone marrow secreted protein; bone marrow stromal cell; cytokine;
 KM cell proliferation; cell differentiation; hematopoiesis; anaemia;
 KM myeloid cell deficiency; lymphoid cell deficiency; myeloid cell;
 KM erythroid progenitor cell; colony stimulating factor; granulocyte;
 monocyte; macrophage; myelo-suppression; megakaryocyte; platelet;
 platelet disorder; thrombocytopenia; hematopoietic stem cell;
 stem cell disorder; aplastic anaemia; bone differentiation;
 KM paroxysmal nocturnal hemoglobinuria; bone growth; cartilage; tendon;
 KM ligament; nerve; wound healing; tissue repair; burn; incision; ulcer;
 KM bone fracture; cartilage damage; artificial joint.

OS Homo sapiens

Key	Location/Qualifiers
FH	1.62
FT	
FT	/note= "signal peptide"

PN W09933979-A2.

PD 08-JUL-1999.

PF 18-DEC-1998; 98WO-US27008.

PR 30-DEC-1997; 97US-0068958.

PR 30-SEP-1998; 98US-0102540.

PA (CHIR) CHIRON CORP.

PI Lin H, Cao L;

XX

DR WPI; 2000-038344/03
DR N-PSDB; AA236238.

PT New isolated human polynucleotide and secreted proteins can induce
PT production of other cytokines in certain cell populations -

PS Claim 2; Page 94; 120pp; English.

CC AMV5362243 represent bone marrow secreted proteins of human bone marrow
 CC stromal cells. The proteins can exhibit cytokine, cell proliferation, or
 CC cell differentiation activity (either inducing or inhibiting). They can
 CC be used to support colony forming cells or factor-dependent cell lines,
 CC to regulate hematopoiesis, and to treat myeloid or lymphoid cell
 CC deficiencies. In addition, they may be used to support the growth and
 CC proliferation of erythroid progenitor cells, and to treat various
 CC anaemias. They can have colony stimulating factor (CSF) activity and can
 CC be used to support the growth and proliferation of myeloid cells such as
 CC granulocytes, monocytes or macrophages, to prevent or treat
 CC myeloid suppression, to support the growth and proliferation of
 CC megakaryocytes and platelets, thereby allowing prevention or treatment
 CC of platelet disorders such as thrombocytopenia, to support the growth
 CC and proliferation of hematopoietic stem cells, either in place of or in
 CC conjunction with platelet transfusions, to treat stem cell disorders,
 CC such as aplastic anaemia and paroxysmal nocturnal hemoglobinuria, or to
 CC repopulate the stem cell compartment after irradiation or chemotherapy.
 CC They can be used for growth or differentiation of bone, cartilage,
 CC tendon, ligament, or nerve tissue, as well as for wound healing and
 CC tissue repair and replacement, and in the treatment of burns, incisions
 CC and ulcers, to induce cartilage and/or bone growth in circumstances
 CC where bone is not normally formed and thus have an application in healing
 CC bone fractures and cartilage damage or defects, prophylactic use in
 CC fracture reduction and also in the improved fixation of artificial
 CC joints.

50 Sequence 167 AA,

Query Match	20.9%	Score 161.5	DB 21	Length 167
Best Local Similarity	35.6%	Pred. No. 8.4e-12		
Matches 36; Conservative	20;	Mismatches 44;	Indels 1;	Gaps 1.

[illegible]

Db 95 sgsIsiaqirttkqIqIdgmVllsvIefclavslsafgc 135

RESULT	3	.
AAW06503		
ID	AAW06503	standard; Protein; 214 AA

AC AAW06503;

DT 06-FEB-1997 (first entry)

DE HTm4 protein

KW Human HTM4 protein; Fc(epsilon)RI receptor; FR: atopic disease; allergy; asthma; atopic dermatitis; allergic rhinitis; hereditary

OS Homo sapiens

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

FT	/note= "Transmembrane domain"
----	-------------------------------

FT /note= "Transmembrane domain"

```
FT
/note= "Transmembrane domain"
```

FT	Domain	173..192
----	--------	----------

XX 25-MAY-2000.
 PD 04-NOV-1999; 99MO-US26234.
 XX 18-NOV-1998; 98US-0195292.
 PR (INCY-) INCYTE PHARM INC.
 PA Walker MG, Volkmut W, Klingler TM;
 PI WPI: 2000-387787/33.
 DR N-PSDB; AAA27130.
 XX New human inflammation-associated polypeptide useful for diagnosis,
 PT prevention and treatment of inflammatory diseases comprises product of
 PT gene coexpressed with e.g. CD16, L-selectin and IP-30
 XX Claim 4; Page 39-40; 43pp; English.

CC Eleven novel inflammation-associated genes have been identified in
 CC CDNA libraries from various tissues. The genes were selected
 CC according to their coexpression with the known inflammation genes,
 CC CD16, L-selectin, Src-like adapter protein, IP-30, superoxide
 CC homoenzyme subunits, alpha-1-antitrypsin, C1q-A, 5-lipoxygenase
 CC activating protein and Src family tyrosine kinase. The novel
 CC polynucleotides may be used in hybridization assays to diagnose a
 CC disease or condition associated with altered expression of the
 CC inflammation genes. Antibodies against the genes may be useful in
 CC compositions for the diagnosis and treatment of such diseases
 CC associated with inflammation including rheumatoid arthritis,
 CC Crohn's disease, multiple sclerosis, AIDS, diabetes mellitus,
 CC asthma and allergy. Additionally the polynucleotides of the
 CC invention may be used for gene therapy. The present sequence is
 CC human inflammation associated protein #6, derived from Incyte
 CC Clone 2349263.
 XX Sequence 225 AA;
 SQ

Query Match 18.0%; Score 139; DB 21; Length 225;
 Best Local Similarity 26.5%; Pred. No. 6.7e-09;
 Matches 41; Conservative 31; Mismatches 73; Indels 10; Gaps 4;
 OY 1 MDSTASHPVFLVPEPTASEYESTELSATFTSTOSPLQKLPARKMKILGTIQLFGIM 60
 1 mtsqpyneilivlpnvl--nfsqaekpeptngqdsikkhhaelkvgtlqilgmm 58
 61 TFSFGVIFLTLKP-YRPPFIPL-SGYPFGSVLFINSAGPLIAVKRTTETL----- 113
 59 vlsigllsaasfsptftgstlinsaypflgpfiflisstistatekrlklvhsnl 118
 OY 114 -GILITLMPFSIIEFLFSILGCHSEDCDEQ 147
 119 vgsilsalselvgfllsvkqaclnpaslqceldk 153
 Db

RESULT 6
 AAY73495
 ID AAY73495 standard; Protein; 242 AA.
 AC AAY73495;
 XX 29-FEB-2000 (first entry)
 DT Human secreted protein clone y33_1 protein sequence SEQ ID NO:212.
 XX Human; secreted protein; immunostimulatory; haemostatic; cytokine;
 KW proliferative; differentiative; chemotactic; chemokinetic; vaccine;
 KW thrombolytic; antiinflammatory; cyostatic; immunosuppressive;
 KW gene therapy.
 XX Homo sapiens.
 OS

XX MO958642-A2.
 PN 18-NOV-1999.
 XX 14-MAY-1999; 99MO-US10843.
 PR 14-MAY-1998; 98US-0085472.
 PR 17-AUG-1998; 98US-0096824.
 PR 11-SEP-1998; 98US-0099843.
 PR 11-SEP-1998; 98US-0099950.
 PR 15-SEP-1998; 98US-0100424.
 PR 29-SEP-1998; 98US-0102329.
 PR 09-OCT-1998; 98US-0103615.
 PR 11-DEC-1998; 98US-0111799.
 PR 14-DEC-1998; 98US-0112159.
 PR 31-DEC-1998; 98US-0114415.
 PR 10-FEB-1999; 99US-0248059.
 PR 06-APR-1999; 99US-0287150.
 PR 13-MAY-1999; 99US-0311021.
 XX

PA (GEM) GENETICS INST INC.
 PI Wong GG, Clark HF, Fechtel K, Agostino MJ;
 DR WPI: 2000-053095/04.
 DR N-PSDB; AA252580.
 XX Novel polynucleotides and proteins having biological activities which
 PT make them suitable for treating, preventing or ameliorating medical
 PT conditions in humans or animals -
 XX Claim 221; Page 718; 730pp; English.

CC The present invention describes human secreted proteins encoded by
 CC polynucleotides obtained from adult testes, foetal brain, adult brain,
 CC brain (foetal and adult), foetal kidney, adult spleen, and adult thymus
 CC CDNA libraries. The polynucleotides and proteins are predicted to have
 CC biological activities which would make them suitable for treating,
 CC preventing or ameliorating medical conditions in humans and animals.
 CC Suggested activities include nutritional activity, cytokine and cell
 CC proliferation/differentiation activity, immune stimulating (e.g. as
 CC vaccines) or suppressing activity, haematopoiesis regulating activity,
 CC tissue growth activity, activin/inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic and thrombolytic activity, receptor/
 CC ligand activity, anti-inflammatory activity, cachectin/tumour invasion
 CC suppressor activity, and tumour inhibition activity. The polynucleotides
 CC are also stated to be useful for gene therapy. Therapeutic compositions
 CC are also presently valuable for veterinary applications. AA252475 to
 CC AA252581 encode human secreted proteins, and AAY73390 to AAY73500
 CC represent human secreted proteins, given in the present invention.
 XX Sequence 242 AA;
 SQ

Query Match 18.0%; Score 139; DB 21; Length 242;
 Best Local Similarity 26.5%; Pred. No. 7.4e-09;
 Matches 41; Conservative 31; Mismatches 73; Indels 10; Gaps 4;
 OY 1 MDSTASHPVFLVPEPTASEYESTELSATFTSTOSPLQKLPARKMKILGTIQLFGIM 60
 1 mtsqpyneilivlpnvl--nfsqaekpeptngqdsikkhhaelkvgtlqilgmm 58
 61 TFSFGVIFLTLKP-YRPPFIPL-SGYPFGSVLFINSAGPLIAVKRTTETL----- 113
 59 vlsigllsaasfsptftgstlinsaypflgpfiflisstistatekrlklvhsnl 118
 OY 114 -GILITLMPFSIIEFLFSILGCHSEDCDEQ 147
 119 vgsilsalselvgfllsvkqaclnpaslqceldk 153
 Db

RESULT 7

AAV15225
 ID AAV15225 standard; protein; 248 AA.
 AC AAV15225;
 XX
 DT 26-OCT-1999 (first entry)
 XX
 DE Human receptor protein (HURP) 4 amino acid sequence.
 XX
 KW receptor; cancer; autoimmune disorder; inflammation;
 KW antagonist; cell surface protein; cell signalling;
 KW antibody; human receptor protein; HURP; reproductive disorder;
 KW developmental disorder; gastrointestinal disorder.
 XX
 OS Homo sapiens.
 XX
 PH Key Location/Qualifiers
 FT Misc-difference 8 /note= "Potential N-glycosylation site"
 FT Misc-difference 20 /note= "Potential N-glycosylation site"
 FT Misc-difference 74 /note= "Potential N-glycosylation site"
 FT Misc-difference 110 /note= "Potential N-glycosylation site"
 FT Misc-difference 22 /note= "Potential CAMP-/cGMP-dependent protein-kinase phosphorylation site"
 FT Misc-difference 22 /note= "Potential casein kinase II-phosphorylation site"
 FT Misc-difference 193 /note= "Potential casein kinase II-phosphorylation site"
 FT Misc-difference 193 /note= "Potential casein kinase II-phosphorylation site"
 FT Misc-difference 36 /note= "Potential protein kinase C-phosphorylation site"
 FT Misc-difference 105 /note= "Potential protein kinase C-phosphorylation site"
 FT Misc-difference 136 /note= "Potential protein kinase C-phosphorylation site"
 FT Misc-difference 177 /note= "Potential protein kinase C-phosphorylation site"
 FT Misc-difference 233 /note= "Potential protein kinase C-phosphorylation site"
 FT Misc-difference 242 /note= "Potential tyrosine kinase phosphorylation site"
 FT Misc-difference 242 /note= "Potential tyrosine kinase phosphorylation site"
 PN WO941375-A2.
 XX
 PD 19-AUG-1999.
 XX
 PF 05-FEB-1999; 99WO-US02572.
 XX
 PR 12-FEB-1998; 98US-0022939.
 XX
 PA (INCYTE) INCYTE PHARM INC.
 XX
 PI Au-Young J, Bandman O, Baughn M, Corley NC, Guegler KU;
 PI Hillman JL, Lal P, Shah P, Tang YT, Yue H;
 DR WPI; 1999-494536/41.
 DR N-PSDB; AA206369.
 XX
 PT New human receptor proteins, used e.g. to treat, prevent and
 PT diagnose gastrointestinal and developmental disorders - and related
 PT nucleic acids, vectors, transformed cells, antibodies, agonists and
 PT antagonists
 PS Claim 1; Page 81-82; 94pp; English.

XX The human receptor protein 4 (HURP-4) has 22% homology with the
 CC rat 19F receptor and 19% homology with human CD20 protein.
 CC HURP-4 is expressed in cancerous, inflamed, hematopoietic/immune
 CC and gastrointestinal tissue. HURP-4 therefore appears to have a role in
 CC some forms of cancer, autoimmune/inflammatory disorders, and
 CC gastrointestinal disorders.
 CC This gives rise to the possibility of using an antagonist or an antibody
 CC of HURP to treat or prevent cancer or autoimmune/inflammatory disorders.
 CC
 SQ Sequence 248 AA;
 Query Match 18.0%; Score 139; DB 20; Length 248;
 Best Local Similarity 26.5%; Pred. NO. 7.6e-09;
 Matches 41; Conservative 31; Mismatches 73; Indels 10; Gaps 4;
 QY 1 MDSSSTANSPVFLVPPETITASEVSTELSAVTFSTQSPQKLFARKMKILGTIOILFGIM 60
 DB 1 mtsqpyneclivpsnvl--nfsqaekpeptngqskhlnhaelkvgtlqlcgmm 58
 QY 61 TFSFGVTFLETLKP-YRPEPFIPL-SGYPPWGSVLFINSGLFIIVAKRKTETL----- 113
 DB 59 vlsigllasasfspntfyvstllnsayftgpfiffllsgslatekrtikllyhssl 118
 QY 114 -GLITLMTFSLIELFISLPSILGCHSEDCDCHQ 147
 DB 119 vgsllsalsalvgfllsvkqatlnpaslqcelak 153
 RESULT 8
 ID AAV91531
 AC AAV91531 standard; protein; 248 AA.
 DT 29-JUN-2000 (first entry)
 XX
 DE Human secreted protein sequence encoded by gene 81 SEQ ID NO:204.
 KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
 KW antiHIV; antiinflammatory; neutropic; neuroprotective; antiasthmatic;
 KW osteopathic; antiarthritic; antibacterial; antidiabetic; antisthma;
 KW antiparasitic; cardiac; gene therapy; cancer; neurological disorder;
 KW immune disease; inflammation; blood disorder; tumour; chromosome 1.
 XX
 OS Homo sapiens.
 XX
 PN WO200006698-A1.
 XX
 PD 10-FEB-2000.
 XX
 PF 29-JUL-1999; 99WO-US17130.
 XX
 PR 30-JUL-1998; 98US-0094657.
 PR 05-AUG-1998; 98US-0095486.
 PR 06-AUG-1998; 98US-0095454.
 PR 06-AUG-1998; 98US-0095455.
 PR 12-AUG-1998; 98US-0096319.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Komatsu S, GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;
 PI Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;
 PI Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenski M;
 DR WPI; 2000-195282/17.
 DR N-PSDB; AAA26426.
 XX
 PT New isolated human genes and the secreted polypeptides they encode,
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders -

PS Claim 11; Page 513-514; 634pp; English.

XX The polynucleotide sequences given in AAB70489 to AAB70490 encode the
CC human secreted proteins given in AAB70489 to AAB70490. The human secreted
CC proteins can have activities based on the tissues and cells they are
CC expressed in. Examples of the activities are: cytostatic;
CC immunosuppressive; anti-HIV; anti-inflammatory; neurotrophic; neuroprotective;
CC anti-angiogenic; osteoprotective; antiarthritic; antibacterial; antidiabetic;
CC antiasthma; antipsoriatic; and cardiac. The polynucleotides and their
CC corresponding secreted proteins are useful for preventing, treating or
CC ameliorating medical conditions, e.g. by protein or gene therapy. Also
CC pathological conditions can be diagnosed by determining the amount of the
CC proteins in a sample or by determining the presence of mutations in the
CC polynucleotides. Specific uses are described for each of the
CC polynucleotides, based on which tissues they are most highly expressed
CC in, and include developing products for the diagnosis or treatment of
CC cancer, tumors, neurodegenerative disorders, developmental abnormalities
CC and foetal deficiencies, blood disorders, diseases of the immune system,
CC autoimmune diseases, hepatic and renal disorders, inflammation,
CC allergies, Alzheimer's and behavioural disorders, schizophrenia,
CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
CC cardiovascular disorders, reproductive disorders, gastrointestinal
CC disorders, respiratory disorders and metabolic disorders. The proteins
CC or polynucleotides can also be used as food additives or preservatives.
CC The proteins are also useful for identifying their binding partners.
CC AAB70489 to AAB70490 and AAB70491 are sequences used in the
CC exemplification of the present invention.

SO Sequence 248 AA;

Query Match 18.0%; Score 139; DB 21; Length 248;
Best Local Similarity 26.5%; Pred. No. 7.6e-09;
Matches 41; Conservative 31; Mismatches 73; Indels 10; Gaps 4;

QY 1 MDSSTAHSPVFLVPPPEITASEYESTELSTSTQSPLOKLFARKKKIIGTIIQILFGIM 60
DB 1 mtsqpyneclivlpnsnvi--nfsgaekpeptngqgskkhhaeklvigtlqilcgm 58
QY 61 TFSFGVIFLTLKP-YRPPPEIFL-SGYPPWGSVLFINSGAFLIAVKKRTETL----- 113
DB 59 vlsigillaaasfsfntqvtstlinsaypfifgffiflsgislsatekrltkilvns1 118
QY 114 -GILITMTFSIIEFLSLPSILGCHSEDCDCQ 147
DB 119 vgsllsalsalvgfllsvkqatlnpsalqceldk 153

RESULT 9

AAB70489 AAB70489 standard; Protein; 248 AA.

XX AAB70489;

DT 04-MAY-2001 (first entry)

DE Human hHAIRBs-iso protein sequence SEQ ID NO:7.

KW Human; hHAIRBs-iso; HAIRBs; HAIRBs isomer; detection;

KW high affinity immunoglobulin epsilon receptor beta subunit.

OS Homo sapiens.

XX CNI269410-A.

PD 11-OCT-2000.

PF 17-MAR-2000; 2000CN-0114959.

PR 17-MAR-2000; 2000CN-0114959.

PA (SREN-) SOUTHERN RES CENT NAT HUMAN GENE GROUP.

XX Xiao H, Liu F, Song H;

DR WPI; 2001-050545/07.

DR N-PSDB; AAF63724.

PT New human immunoglobulin receptor subunit protein and its nucleic acid

PS Claim 4; Page 17; 22pp; Chinese.

XX The present invention describes a human high affinity immunoglobulin
CC epsilon receptor beta subunit isomer, designated hHAIRBs-iso.
CC hHAIRBs-iso is isolated from in human pheochromocytoma. The present
CC invention also describes methods for the preparation and detection of
CC hHAIRBs-iso protein and nucleotide sequences. The present sequence
CC represents the human hHAIRBs-iso protein, as given in the present
CC invention.

SO Sequence 248 AA;

Query Match 18.0%; Score 139; DB 22; Length 248;
Best Local Similarity 26.5%; Pred. No. 7.6e-09;
Matches 41; Conservative 31; Mismatches 73; Indels 10; Gaps 4;

QY 1 MDSSTAHSPVFLVPPPEITASEYESTELSTSTQSPLOKLFARKKKIIGTIIQILFGIM 60
DB 1 mtsqpyneclivlpnsnvi--nfsgaekpeptngqgskkhhaeklvigtlqilcgm 58
QY 61 TFSFGVIFLTLKP-YRPPPEIFL-SGYPPWGSVLFINSGAFLIAVKKRTETL----- 113
DB 59 vlsigillaaasfsfntqvtstlinsaypfifgffiflsgislsatekrltkilvns1 118
QY 114 -GILITMTFSIIEFLSLPSILGCHSEDCDCQ 147
DB 119 vgsllsalsalvgfllsvkqatlnpsalqceldk 153

RESULT 10

AAB70489 AAB70489 standard; Protein; 250 AA.

XX AAB70489;

DT 08-DEC-1999 (first entry)

DE Human breast tumour-associated protein 50.

KW Expressed sequence tag; EST; human; breast; cancer; cytostatic;

KW medicaments; gene therapy; treatment; fat metabolism.

OS Homo sapiens.

PN DE19813835-A1.

PD 23-SEP-1999.

PF 20-MAR-1998; 98DE-1013835.

PR 20-MAR-1998; 98DE-1013835.

PA (META-) METAGEN GES GENOMFORSCHUNG MBH.

PI Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;

DR WPI; 1999-528979/45.

DR N-PSDB; AAZ33586.

PT Human nucleic acid sequences and protein products from normal breast

PS tissue, useful for breast cancer therapy -
Claim 28; 180; 206pp; German.

XX This invention describes novel human nucleic acid sequences from normal
 CC breast tissue which have cytostatic activity. The nucleic acid sequences
 CC can be used to produce and isolate full-length gene sequences. They can
 CC be used to express proteins, which can be used as tools to find an
 CC activity against breast cancer. The sequences can be used in sense or
 CC antisense form. They are especially useful for medications for gene
 CC therapy to treat breast cancer and for treating illnesses associated
 CC with fat metabolism. AAY48456-Y48539 represent protein fragments encoded
 CC by the expressed sequence tags described in the method of the invention.
 XX
 SQ Sequence 250 AA;
 Query Match 18.0%; Score 139; DB 20: Length 250;
 Best Local Similarity 26.5%; Pred. No. 7.7e-09;
 Matches 41; Conservative 31; Mismatches 73; Indels 10; Gaps 4;
 QY 1 MDSSTASPVLPVPPETASEYESTELATFTFSTPQKLFARKMKILGTIOILFGIM 60
 Db 26 mtsgpvpnetivlpnsvi--nfsqaekpeptngqdsikkhlaelkvigtllqgm 83
 Y 61 TFSFGVIFLTLKP-YRPFPIFL-SGYPWGSVLFNSGAFILAVKRKTETL----- 113
 Db 84 vlsigllasasfnpftqvstllnsaypfifgpfllsgslstakekrlkllvhs1 143
 QY 114 -GLITLMTFSIIEFLISLPSILGCHSEDDCQ 147
 Db 144 vgsllsalsalvgfllsvkqatlnpslqcelck 178
 RESULT 11
 AAB58419
 ID AAB58419 standard; Protein: 257 AA.
 AC AAB58419;
 DT 14-MAR-2001 (first entry)
 XX
 DE Lung cancer associated polypeptide sequence SEQ ID 757.
 XX
 KW Human; lung cancer associated protein; neuroprotective; cytostatic;
 KW cardioactive; immunomodulatory; muscular active; vulnery;
 KW gastrointestinal; nephrotropic; antinefective; gynecological;
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
 KW proliferative disorder; wound healing; infectious disease.
 XX
 OS Homo sapiens.
 WO20005180-A2.
 PD 21-SEP-2000.
 PF 08-MAR-2000; 2000WO-US05918.
 PR 12-MAR-1999; 99US-0124270.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.
 PI Ruben SM;
 DR WPI: 2000-587514/55.
 DR N-PSDB; AAF18295.
 XX
 PT Lung cancer associated gene sequences, referred to as lung cancer
 PT antigens, useful for treatment, prevention, and diagnosis of disorders
 PT such as lung cancer -
 XX
 PS Claim 11: Page 1274-1275; 1425pp; English.
 CC Polynucleotide sequences AAF1982 - AAF18424 encode human lung cancer
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer

CC associated proteins and polynucleotide sequences, their agonists, and
 CC antagonists may have neuroprotective; cytostatic; cardioactive;
 CC immunomodulatory; muscular active general; vulnery; gastrointestinal
 CC general; nephrotropic; antinefective; gynecological; or antibacterial
 CC activity. The invention also includes antibodies specific for the
 CC protein or polynucleotide sequences. The lung cancer associated
 CC polynucleotide sequences may be used for detection of lung cancer,
 CC chromosome identification, as chromosome markers, and for numerous other
 CC diagnostic or research purposes. The proteins may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders. The proteins may also be used in the treatment of wounds and
 CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
 CC peptide AAB58549 are used in the course of the invention for the
 CC identification and characterisation of the polynucleotide and protein
 CC sequences.
 XX
 SQ Sequence 257 AA;
 Query Match 18.0%; Score 139; DB 21: Length 257;
 Best Local Similarity 26.5%; Pred. No. 7.9e-09;
 Matches 41; Conservative 31; Mismatches 73; Indels 10; Gaps 4;
 QY 1 MDSSTASPVLPVPPETASEYESTELATFTFSTPQKLFARKMKILGTIOILFGIM 60
 Db 26 mtsgpvpnetivlpnsvi--nfsqaekpeptngqdsikkhlaelkvigtllqgm 83
 Y 61 TFSFGVIFLTLKP-YRPFPIFL-SGYPWGSVLFNSGAFILAVKRKTETL----- 113
 Db 84 vlsigllasasfnpftqvstllnsaypfifgpfllsgslstakekrlkllvhs1 143
 QY 114 -GLITLMTFSIIEFLISLPSILGCHSEDDCQ 147
 Db 144 vgsllsalsalvgfllsvkqatlnpslqcelck 178
 RESULT 12
 AAY91680
 ID AAY91680 standard; Protein: 273 AA.
 AC AAY91680;
 DT 29-JUN-2000 (first entry)
 XX
 DE Human secreted protein sequence encoded by gene 81 SEQ ID NO:353.
 XX
 KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
 KW antiHIV; antiinflammatory; nootropic; neuroprotective; antiallergic;
 KW osteopathic; antiarthritic; antibacterial; antidiabetic; antisthma;
 KW antiporiatic; cardiant; gene therapy; cancer; neurological disorder;
 KW immune disease; inflammation; blood disorder; tumour; Chromosome 1.
 XX
 OS Homo sapiens.
 PN WO200006698-A1.
 PD 10-FEB-2000.
 PF 29-JUL-1999; 99WO-US17130.
 PR 30-JUL-1998; 98US-0094657.
 PR 05-AUG-1998; 98US-0095486.
 PR 06-AUG-1998; 98US-0095454.
 PR 06-AUG-1998; 98US-0095455.
 PR 12-AUG-1998; 98US-0096319.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Komatsoulis GA, Rosen SM, Ruben SM, Duan R, Moore PA, Shi Y;
 PI Lafleur D, Wei Y, Nl J, Florence KA, Young PE, Brewer LA;
 PI Sopet DR, Endress GA, Ebner R, Olsen HS, Mucenski M;


```

XX Key Location/Qualifiers
FH Modified-site 125
FT /note= "potential protein kinase C phosphorylation
FT site"
FT Modified-site 137
FT /note= "potential N-glycosylation site"
FT Modified-site 192
FT /note= "potential casein kinase II phosphorylation
FT site"
XX US5871930-A.
XX 16-FEB-1999.
XX 21-AUG-1997: 97US-0916902.
XX 21-AUG-1997: 97US-0916902.
XX (INCY-) INCYTE PHARM INC.
XX Bandman O, Corley NC, Lal P;
XX MPI; 1999-166628/14.
XX N-PSDB; AAX15104.
XX New polynucleotide coding for high affinity immunoglobulin E
XX receptor-like protein - and its complement, useful for diagnosis,
XX prevention and antisense therapy of inflammatory responses
XX Claim 1; Fig 1A-D; 30pp; English.
XX
XX The present sequence represents a high affinity immunoglobulin E
XX receptor-like protein (IGERB). The present sequence was first
XX identified in Incyte clone 927955 for the brain cDNA library
XX BRAINOT04. Polynucleotides complementary to the IGERB cDNA can
XX be used as probes to IGERB gene expression. IGERB proteins and
XX nucleotides may be used for the diagnosis of inflammatory
XX responses associated with expression of IGERB, e.g. AIDS,
XX Addison's disease, atherosclerosis, bronchitis, ulcerative colitis,
XX diabetes mellitus, emphysema, gout, Graves's disease, osteoporosis,
XX rheumatoid arthritis, Sjogren's syndrome and complications of cancer.
XX The sequences may also be useful in assays that detect activation or
XX induction of various cancers.
XX
XX Sequence 239 AA:
SQ
Query Match 17.4%; Score 134; DB 20; Length 239;
Best Local Similarity 31.4%; Pred. No. 2.9e-08;
Matches 38; Conservative 20; Mismatches 47; Indels 16; Gaps 3;
QY 39 LQKLFAR-KKKITGTOILFGIMTFESGCVIFLTLKPYRPFPIFLSGYFPGSVLEFIN 97
DB 54 Iqekfllkgepkvlgvqjllalmslsmgltmmcasnlygsnpisyyigtllgswmfll 113
QY 98 SGAFLLAVKRKTE-----TLGILIT-----LMFSSIELFLSPFSILGCHSED 142
DB 114 sgslslaaagrlrttkglvrgslgmnlssvlaasgllintslsfsthyphycnygnsmn 173
QY 143 C 143
DB 174 c 174
RESULT 15
AAY50174
ID AAY50174 standard; Protein; 239 AA.
XX
XX AAY50174;
XX
XX 31-JAN-2000 (first entry)
XX

```

```

DE Human high affinity IGE receptor-like protein (IGERB).
XX
XX IGERB; high affinity; immunoglobulin E; IGE; receptor; allergy;
XX inflammation; B lymphocyte; homology; high affinity IGE receptor; FCRI;
XX mast cell; basophil; histamine; protease; cytoplasmic granule; synthesis;
XX effector; prostaglandin; leukotriene; cytokine; antigen; antibody;
XX identification; agonist; leukotriene; expression; activity; diagnosis;
XX therapy; inflammatory disorder; multiple sclerosis; osteoarthritis;
XX asthma; cancer; side effect; complication.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Modified-site 125
FT /note= "Phosphorylated by protein kinase C"
FT Modified-site 137
FT /note= "Glycosylated"
FT Modified-site 192
FT /note= "Phosphorylated by casein kinase II"
XX US5977072-A.
XX 02-NOV-1999.
XX 15-DEC-1998: 98US-0213389.
XX 21-AUG-1997: 97US-0916902.
XX (INCY-) INCYTE PHARM INC.
XX Bandman O, Corley NC, Lal P;
XX MPI; 2000-012123/01.
XX N-PSDB; AAZ32842.
XX High affinity immunoglobulin E receptor-like protein useful for
XX stimulating allergic and immune responses -
XX Claim 1; Fig 1; 29pp; English.
XX
XX This sequence represents human high affinity immunoglobulin E (IGE)
XX receptor-like protein (IGERB). Nucleic acids encoding IGERB were
XX initially identified in a brain cDNA library, this sequence being a
XX consensus. An allergic response is initiated by release of IGE
XX from B lymphocytes. The IGE molecules then bind to the high affinity
XX IGE receptor (FCRI) present on mast cells and basophils, which triggers
XX the release of histamine and proteases from cytoplasmic granules and
XX leads to the synthesis of effectors of the allergic and inflammatory
XX response, such as prostaglandins, leukotrienes and cytokines. As IGERB
XX binds IGE, it may be administered to stimulate allergic and immune
XX responses in patients in whom IGERB is under expressed or inactive and
XX to supplement the patients own production of the protein. IGERB may also
XX be used as an antigen for the production of antibodies and to identify
XX candidate agonists and antagonists of IGERB expression and activity.
XX The antibodies may also be used in diagnosis. Antibodies and antagonists
XX may be administered to downregulate IGERB activity and reduce the
XX potency of inflammatory and allergic responses. They may be used in this
XX way to treat inflammatory disorders such as multiple sclerosis,
XX osteoarthritis, asthma and some complications of cancer. Conversely, the
XX agonists may be used to enhance immune responses.
XX
XX Sequence 239 AA:
SQ
Query Match 17.4%; Score 134; DB 21; Length 239;
Best Local Similarity 31.4%; Pred. No. 2.9e-08;
Matches 38; Conservative 20; Mismatches 47; Indels 16; Gaps 3;
QY 39 LQKLFAR-KKKITGTOILFGIMTFESGCVIFLTLKPYRPFPIFLSGYFPGSVLEFIN 97
DB 54 Iqekfllkgepkvlgvqjllalmslsmgltmmcasnlygsnpisyyigtllgswmfll 113
QY 98 SGAFLLAVKRKTE-----TLGILIT-----LMFSSIELFLSPFSILGCHSED 142

```

Thu Oct 11 10:24:08 2001

us-09-735-712-8.rag

Page 10

Db 114 sgs1s1aag1rttckgylvrsglgnmtssvlaasgllntfslafysfhpypcnyygnsn 173
QY 143 C 143
Db 174 c 174

Search completed: October 7, 2001, 04:39:26
Job time: 1911 sec


```

QY      1 MDSTAH-SPVFLVPPPEITASEYESTELSATTFTSQSPLOKLFARKMILGTIQLLFGI 55
      : : : : : | : : : : :
Db-    11 LGSASAHGTPGSETGPDELNTSYH-----PINSPPDYK---AKLOVLGAIQILNAA 60C

```

[illegible]

```
DB      120 NSFQNMNIASATIALVGTAFSLNIAVNIGLSKSSSSSPDLG 163

RESULT      3
US-08-916-902A-3
; Sequence 3, Application US/08916902A
; Patent No. 5871930
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
; TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: Herewith
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0371 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 214 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 561639
; US-08-916-902A-3

Query Match          19.5%; Score 150.5; DB 2; Length 214;
Best Local Similarity 31.1%; Pred. No. 2,le-10;
Matches    51; Conservative   22; Mismatches   64; Indels   27; Gaps     8.
```

```

APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/213,389
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/916,902
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0371 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO.: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 214 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 561639
US-09-213-389-3

Query Match          19.5%, Score 150.5; DB 2; Length 214;
Best Local Similarity 31.1%; Pred. No. 2.1e-10;
Matches 51; Conservative 22; Mismatches 64; Indels 27; Gaps 8;

OY      1 MDSIAH-SPTFLVPPETASEYSTELSATTFSTQSPLOKLFARKKKILGTLIGCI 59
       : ::::: : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      11 LGSASAHGTPESENGPBEHLNLSVYH-----PIHGSDPYOK--AKLVGLAIDILNA 60
OY      60 MTFESGVLEFLTKLPY----PRPFIFLGGYPFGSVLEFINSGAFLLAVKRKTRET--- 112
       | : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      61 MILAGV-FLGSLQPYHFQKHFFFTYTGTGPINGAVFPCCSSGLTSLVYAGIKPRTMIQ 119
OY      113 --LGLITLMTFSII-ELFISLPF-----SILGHSDEDCDEQC 148
       | : | : : : : : : : : : : : : : : : : : : : : : : : :
Db      120 NSFGNMASATIALVGTAFLSLNLIAVNIOQLRSCHSSSESPLDC 163

RESULT        6
US-08-916-902A-1
: Sequence 1, Application US/08916902A
: Patent No. 5871930
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lai, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESS: Incyte Pharmaceuticals, Inc.
```

STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/916,902A
FILING DATE: Herewith
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0371 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 239 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT04
CLONE: 927955
US-08-916-902A-1

Query Match 17.4%; Score 134; DB 2; Length 239;
Best Local Similarity 31.4%; Pred. No. 2.3e-08;
Matches 38; Conservative 20; Mismatches 47; Indels 16; Gaps 3;

QY 39 LQKLFAR-KKKILGCTIOLFGIMTFSGVIFLTLKPYRPPPIFLSGIPFGSVLFIN 97
DB 54 LQEFELGEPKVLGVOLILALMSLSMGITMCMASNTGSPISYIGYIMGSVMFII 113
QY 98 SGAFLIAVKRRTTE-----TLGILIT-----LMTFSIILFLSLPFSILGCHSED 142
114 SGLSISLAGIRTTKGLVRSGLGMNITSSVLAASGILINTFSLAFYRHHPCNTYNGNSNN 173

QY 143 C 143
DB 174 C 174

RESULT 7
US-09-213-389-1
Sequence 1, Application US/09213389
Patent No. 5977072
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/213,389
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/916,902
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0371 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 239 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT04
CLONE: 927955
US-09-213-389-1

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/213,389
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/916,902
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0371 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 239 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT04
CLONE: 927955
US-09-213-389-1

Query Match 17.4%; Score 134; DB 2; Length 239;
Best Local Similarity 31.4%; Pred. No. 2.3e-08;
Matches 38; Conservative 20; Mismatches 47; Indels 16; Gaps 3;

QY 39 LQKLFAR-KKKILGCTIOLFGIMTFSGVIFLTLKPYRPPPIFLSGIPFGSVLFIN 97
DB 54 LQEFELGEPKVLGVOLILALMSLSMGITMCMASNTGSPISYIGYIMGSVMFII 113
QY 98 SGAFLIAVKRRTTE-----TLGILIT-----LMTFSIILFLSLPFSILGCHSED 142
114 SGLSISLAGIRTTKGLVRSGLGMNITSSVLAASGILINTFSLAFYRHHPCNTYNGNSNN 173

QY 143 C 143
DB 174 C 174

RESULT 8
US-07-869-933-29
Sequence 29, Application US/07869933
Patent No. 5770396
GENERAL INFORMATION:
APPLICANT: Kinet, Jean-Pierre
TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
NUMBER OF SEQUENCES: 34.
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/869,933
FILING DATE: 19920416

Db 107 LFVLSGFLSMERKNTLYLVKSGANIVSSIAAGLGIALLLNLS 153

```

; LENGTH: 243 amino ac
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```

Db 107 LEVLGGLIMSEKNTLYLVGSLGANIVSSIAAGLGIALLLNTLS 153

CORRESPONDENCE ADDRESS:

Query Match	16.0%;	Score 123.5;	DB 1;	Length 243;
Best Local Similarity	29.0%;	Pred. No. 4,2e-07;		
Matches 31;	Conservative 16;	Mismatches 41;	Indels 19;	Gaps 2;

J. Immunol. 144:4368-4394, 1988

A>Title: Cloning of a complementary DNA encoding a new mouse B lymphocyte differentiation antigen

A:Reference number: A30558; MUID:89067519

A:Accession: A30558

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-291 <TED>

C:Genetics:

A:Map position: 19

C:Superfamily: B-cell surface antigen CD20

C:Keywords: membrane protein; phosphoprotein; surface antigen

Query Match 16.9%; Score 172; DB 2; Length 291;
Best Local Similarity 29.4%; Pred. No./e-09;
Matches 53; Conservative 29; Mismatches 64; Indels 34; Gaps 7;

OY 32 TESTOSPLQKLFARKMKILGTIIQLIFGIMTFSGVIFLTLLKPYRPFPIELS-GYPEF 90
| | | | | : | | : | | : | | : | | : | | : | |
27 TSSLVGPTQSFFMRRESALAGAVQIMNGLFIHTTLGL-----LMIPGVFAPICLSTWYPIM 82
| | | | | : | | : | | : | | : | | : | | : | |
OY 91 GSYLETNSGAFELIAVRKRRTETLIISLR-IMNLSAIALRAIGIIL-----LTFEGFIID 142
| : | | | : | | : | | : | | : | | : | | : | |
DB 83 GGIMTYITSGSLLAAMAAEKTSRKRSYLKAKVIMSSLSFLPAISGIILIMDIINMTISHPELK 142
| : | | | : | | : | | : | | : | | : | | : | |
OY 143 QN-----YIC---GYSHONS-----QCKATVLEFLGITLIMTSIIEFLFS 181
| | | | | : | | : | | : | | : | | : | | : | |
DB 143 MRRELIIQTSPKPYVDIYDCERPSNSEKNSPSTOYCNSIQSVFLGISAMLISAFOKLVT 202
| | | | | : | | : | | : | | : | | : | | : | |

RESULT 3

B-cell surface antigen CD20 - human

N:Alternate names: B-lymphocyte antigen CD20; B1

C:Species: Homo sapiens (man)

C:Date: 08-Jun-1989 #sequence.revision 02-Aug-1996 #text_change 22-Jun-1999

C:Accession: A30586; J10042; A27400; S00387

R:Redder, T.F.; Klejman, G.; Schlossman, S.F.; Saito, H.

J. Immunol. 142, 2560-2568, 1989

A>Title: Structure of the gene encoding the human B lymphocyte differentiation antigen C

A:Reference number: A30586; MUID:89176281

A:Accession: A30586

A:Molecule type: DNA

A:Residues: 1-297 <TED>

A>Note: The authors translated the codon ATG for residue 148 as His, TCA for residue 289 and GAC for residue 294 as Ser

Stamenkovic, I.; Seed, B.

Exp. Med. 167, 1975-1980, 1988

Title: Analysis of two cDNA clones encoding the B lymphocyte antigen CD20 (B1.Bp35), a

A:Reference number: J10042; MUID:88258386

A:Accession: J10042

A:Molecule type: mRNA

A:Residues: 1-297 <STA>

A:Cross-references: GB:X12530; NID:g29773; PIDN:CAA31046.1; PID:g29774

R:Redder, T.F.; Streuli, M.; Schlossman, S.F.; Saito, H.

Proc. Natl. Acad. Sci. U.S.A. 85, 208-212, 1988

A>Title: Isolation and structure of a cDNA encoding the B1 (CD20) cell-surface antigen C

A:Reference number: A27400; MUID:88124792

A:Accession: A27400

A:Molecule type: mRNA

A:Residues: 1-297 <TE2>

A:Cross-references: GB:M27394; GB:J03574; NID:g179307; PIDN:AAA35581.1; PID:g179308

R:Reinfeld, D.A.; Brown, J.P.; Valentine, M.A.; Clark, E.A.; Ledbetter, J.A.

EMBO J. 7, 711-717, 1988

A>Title: Molecular cloning of the human B cell CD20 receptor predicts a hydrophobic prot

A:Reference number: S00387; MUID:88283639

A:Accession: S00387

A:Molecule type: mRNA

A:Residues: 1-112, 'L', 14-297 <EIN>

A:Cross-references: EMBL:X07203; NID:g29775; PIDN:CAA30179.1; PID:g29776

A:Experimental source: David cells

C:Comment: This protein appears not to be glycosylated. Isoforms of varying molecular we

C:Comment: This protein plays an important role in B cell activation.

C.Genetics:
A.Gene: GDB:A.CD20
A.Cross-references: GDB:119761; OMIM:112210
A.Map position: 11q12-11q13.1
C.Superfamily: B-cell surface antigen CD20
C.Keywords: B-cell; phosphoprotein; transmembrane protein
F:1-51/Domain: Intracellular #status predicted <CYT>
F:52-103/Domain: transmembrane #status predicted <TM12>
F:104-116/Domain: intracelllular #status predicted <CYT>
F:117-141/Domain: transmembrane #status predicted <TM3>
F:142-185/Domain: extracelllular #status predicted <EX1>
F:186-212/Domain: transmembrane #status predicted <TM4>
F:213-297/Domain: intracelllular #status predicted <CYT>
F:167-183/Distulfide bonds: #status predicted

Query Match 16.6%; Score 168; DB 1; Length 297;
Best Local Similarity 30.3%; Pred.No. 1.7e+08;
Matches 54; Conservative 27; Mismatches 63; Indels 34; Gaps 7;

OY 34 STQSPLOKIFARKMKILGTIQILEGIMTFPSFGVIFLTLKPFR-FPIFLSGYPWGS 92
| | : | : | : | : | : | : | : | : | : | : | : |
Db 36 SLVGPTQSFPRESRTKLGAIVMNGLFHIALGL-----LMIPAGIADICVTWPPLMG 91

OY 93 VLEINSGALLAVKRKTETILLISRMNLSALANAIGII-----T 136
:: || : | : | : | : | : | : | : | : | : | : | : |
Db 92 IMYITGISLLATEKNRSRCLVKGRMINMSLTFAISGMILTMDLINIKISHLKMES 151

OY 137 FGFILDON-YIGGY-----SHONS----CKAAVTVEFLGITLMTFSII-ELEFI 181
|| : | : | : | : | : | : | : | : | : | : | : |
Db 152 LMFIRAHPTYINIYCEPANPSEKNSPSTQVCYSIQSLFTLGTSMLLPARPEQLVIA 209

RESULT 4
A31231
high-affinity TgE receptor beta chain - rat
C.Species: Rattus norvegicus (Norway rat)
C.Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 05-Nov-1999
C.Accession: A31231
R.Kinet, J.P.; Blank, U.; Ra, C.; White, K.; Metzger, H.; Kochan, J.
Proc. Natl. Acad. Sci. U.S.A. 85, 6483-6487, 1988
A.Title: Isolation and characterization of cDNAs coding for the beta-subunit of the h
A.Reference number: A31231; MUID:88320465
A.Accession: A31231
A.Molecule type: mRNA
A.Residues: 1-243 <kin>
A.Cross-references: GB:M2923; GB:J03845; NID:g204116; PIDN:AAA4149.1; PTD:g204117
C.Keywords: immunoglobulin receptor; transmembrane protein

Query Match 15.7%; Score 159; DB 2; Length 243;
Best Local Similarity 28.0%; Pred.No. 9.9e+08;
Matches 45; Conservative 29; Mismatches 71; Indels 16; Gaps 3;

OY 36 QSPLOLRARRKKILGTIOILEGIMTFPSFGVIFLTLKP--YPRPFIFLSGGYPFWGSV 93
| : :: : | : | : | : | : | : | : |
Db 47 DQTWSFKKELEFGVGVVLGLCLCFGVYCSTLDSDFEDEVLLVRAGVPFMGAV 106

OY 94 LEFINSGAFLIAVRRKTTEFLILSRIMNLASRLRIAGIIITGFDLDQNYICYGHQN 153
|| : | : | : | : | : | : | : | : | : | : | : |
Db 107 LVVISGFLSINSERNTLYLVKGSIKANIVSSIANGLATAILLNLSNNSAVM----- 159

OY 154 SOCKAVT-----VLFGLITLTMTFSIIIEFLISPESIL 187
: || : | : | : | : | : | : | : | : | : | : | : |
Db 160 NYCKDITEDDGCVTVSFITELVLMFLTLILAFCSAVLIIT 200

RESULT 5
B34342
IGE Fc receptor beta chain - mouse
C.Species: Mus musculus (house mouse)
C>Date: 15-Jun-1990 #sequence_revision 15-Jun-1990 #text_change 05-Nov-1999
C.Accession: B34342

R:Ra, C.; Jouvin, M.H.E.; Kinet, J.P.
 J. Biol. Chem. 264, 15323-15327, 1989
 A:Title: Complete structure of the mouse mast cell receptor for IgE (re-epsilon-RI) and
 A:Reference number: A34342; MUID:89359361
 A:Accession: B34342
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-235 <RAC>
 A:Cross-references: GB:J05019; NID:g193238; PIDN:AAA37601.1; PID:g309225
 C:Keywords: Immunoglobulin receptor; Transmembrane protein

Query Match 15.6%; Score 158.5; DB 2; Length 235;
 Best Local Similarity 30.5%; Pred. No. 1.1e-07;
 Matches 60; Conservative 29; Mismatches 81; Indels 27; Gaps 9;

QY 1 MDSSTAHPVFLVFP-PEITASEYESTELSATPTSTQSPLO-KLFARK-MKLTGTIQL 56
 1 MDREN-RSRADLALPNQESSAPDILLEASPAKAPKQMTFTFLKELEFGATQIL 59
 DB 1 FGIITFEFGYIFLTLKPR-----PRPFIFLSGYFPGSVLFINGATLAVKRTTE 111
 57 FGIITFEFGYIFLTLKPR-----PRPFIFLSGYFPGSVLFINGATLAVKRTTE 111
 60 VGLICCFEFGYIFLTLKPR-----PRPFIFLSGYFPGSVLFINGATLAVKRTTE 116
 QY 112 TLITSRIMMLSLRAIAGIILTFEGFILDONVIGCYSHONSCKAVT-----VLFL 164
 117 YLVKSGISGANIVSSIAAGTGMIL-----NLNNFATMNN-CKNVTEDDGCFAVAF 169
 DB 165 GILITMTFSITELFIS 181
 170 TELVLMMLFTILAFCS 186

RESULT 6
 A42806
 IGE Fc receptor beta chain - human
 C:Species: Homo sapiens (man)
 C:Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 08-Oct-1999
 C:Accession: A42806; S21154
 R:Kuester, H.; Zhang, L.; Brini, A.T.; MacGlashan, D.W.J.; Kinet, J.P.
 J. Biol. Chem. 267, 12782-12787, 1992
 A:Title: The gene and cDNA for the human high affinity immunoglobulin E receptor beta ch
 A:Reference number: A42806; MUID:92316966
 A:Accession: A42806
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-244 <RUE>
 A:Cross-references: GB:M89796; NID:g337417; PIDN:AAA60269.1; PID:g337418
 R:Maekawa, K.; Imagawa, N.; Tanaka, Y.; Harada, S.
 S Lett. 302, 161-165, 1992
 A:Title: Determination of the sequence coding for the beta subunit of the human high-aff
 A:Reference number: S21154; MUID:92339505
 A:Accession: S21154
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-244 <MAE>
 A:Cross-references: GB:D10583; NID:g219881; PIDN:BA01440.1; PID:d1001914; PID:g219882
 C:Genetics:
 A:Introns: 19/2; 62/3; 107/3; 126/3; 179/3; 212/3
 C:Keywords: Immunoglobulin receptor; Transmembrane protein

Query Match 12.2%; Score 123.5; DB 2; Length 244;
 Best Local Similarity 25.6%; Pred. No. 0.00022;
 Matches 50; Conservative 33; Mismatches 93; Indels 19; Gaps 7;
 QY 9 PVLVPEPPETASEYESTELSATPTSTQSPLO-KLFARKMTLIGIQLFGIMTFSG 65
 DB 21 PABEVL--EISPEVSSGRLLKS--ASSPLHLWLTVLYKKEOEGLVQTQLITLWICICFG 76
 QY 66 VIFLFTLLKPRPPPL-FLSGYFPGSVLFINGATLAVKRTTELTILSRIMNL 123
 DB 77 TVVCSVLDSHIEGDISSPKAGYFPGWGAIFFSISGLSISERRNATVIVRSLSGNTA 136

QY 124 SALRAINGIILTRGFILDONVIGCYSHONS-----QKAVTVLFLGILITLMTF-SILE 177
 DB 137 SSIAAGTGITLT-----IINKKSLAYIHHSCOKFEFFKCFMASFTIYVMMFLITLG 192
 QY 178 LFIISPFISILGCHSE 192
 DB 193 LGSNVSILICGAGE 207

RESULT 7
 C86844
 hypothetical protein ysbC [imported] - Lactococcus lactis subsp. lactis (strain IL140
 C:Species: Lactococcus lactis subsp. lactis
 C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 23-Mar-2001
 C:Accession: C86844
 R:Bohlin, A.; Winkler, P.; Mager, S.; Jallion, O.; Malarre, K.; Weissenbach, J.; Eh
 Genome Res. in press, 2001
 A:Title: The complete genome sequence of the lactic acid bacterium.
 A:Reference number: A86825
 A:Accession: C86844
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-307 <STO>
 A:Cross-references: GB:AE005176; NID:g12724776; PIDN:AAK05853.1; GSPDB:GN00146
 C:Experimental source: strain IL1403
 C:Genetics:
 A:Gene: ysbC

Query Match 10.1%; Score 103; DB 2; Length 307;
 Best Local Similarity 31.5%; Pred. No. 0.023;
 Matches 45; Conservative 22; Mismatches 50; Indels 26; Gaps 10;

QY 34 STQSPLOKLFARKKILGTIQLFGIMTFSGYIFL--FTLLKPRPP-FFPLSGYFPG 90
 DB 16 ANQNPIN--ADLRKIVGSPFLASGISNF-VGSIFGIITLVTSGTLFSPQFVGSHPAW 71
 QY 91 --GSYFINSAGFLAVKRTTELTILSRIMMLSLRAIAGIILTRGFILDONVIG 147
 DB 72 IWIGGV--GGIFL-----TSNVLLFPRLGAVQTVIPLIGQILM--GTLIDS--F 116
 QY 148 GYSHONSCKAVTVLFGILITL 170
 DB 117 GWFHA-WQLPMTLMRFLGIITL 138

RESULT 8
 JH0751
 IGE receptor beta chain - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
 C:Accession: JH0751
 R:Bieber, T.; de la Salle, H.; Wollenberg, A.; Hakimi, J.; Chizzonite, R.; Ring, J.;
 J. Exp. Med. 175, 1285-1290, 1992
 A:Title: Human epidermal langerhans cells express the high affinity receptor for immu
 A:Reference number: JH0751; MUID:92235608
 A:Accession: JH0751
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-152 <BIE>
 A:Cross-references: GB:X66187; NID:g396463; PIDN:CAA6955.1; PID:g396464

Query Match 9.9%; Score 100; DB 2; Length 152;
 Best Local Similarity 27.6%; Pred. No. 0.023;
 Matches 32; Conservative 21; Mismatches 53; Indels 10; Gaps 3;
 QY 83 FLSGYFPGSVLFINGATLAVKRTTELTILSRIMMLSLRAIAGIILTRGFILD 142
 DB 26 FKAGYFPGWGAIFFSISGLSISERRNATVIVRSLSGANTASISAGTGITLT--IIN 81
 QY 143 QNVICGSHONS-----QKAVTVLFLGILITLMTF-SILELFIISPFISILGCHSE 192

Db 82 LKSLAYIHHSCKQFETKCFMASFSTETIVVMLFTLILGCSAVSLTICGAGE 137

RESULT 9

EB1306
probable transmembrane transport protein Cj1040c [imported] - Campylobacter jejuni (str
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 31-Mar-2000
C:Accession: EB1306
R:Parkhill, J., Wren, B.W., Mungall, K., Ketley, J.M., Churcher, C., Basham, D., Chilling
C.W., Quail, M., Rajandream, M.A., Rutherford, K.M., VanVleet, A., Whitehead, S., Barre
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyf
A:Reference number: AB1250; MWID:20150912
A:Accession: EB1306
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-387 <PAR>
A:Cross-references: GB:AL139077; GB:AL111168; NID:96968444; PIDN:CAB73296.1; PID:9696844
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: Cj1040c

Query Match
Best Local Similarity 25.0%; Score 96; DB 2; Length 387;
Matches 42; Conservative 31; Mismatches 65; Indels 30; Gaps 9;

Db 14 FPEBITSESESEELSTSTSTOSPLKPKKKKILGTQI--LFQIMFSGVITLFT 71

Db 233 FDEESTSMVLFQVLAAPVSLFGL--LLGKLNQNLHTFYIAGLCMYVIAFGMLFI 290

Db 72 LKRPYFPFPI--FLSGYFPGSVLFINSAGFLIAVKRKTETLLISRLMNLALRA 128

Db 291 -----SKISIIISAFIMGFP-WGVF-----GIALFLIAQSSN-----AQIAANLSALAQ 335

Db 129 IAGIILTFEFLIDQNYICGYSR--ONSOKAVTFLGILITLMTF 173

Db 336 GFGYLIAGQ-----QWIIIGFLHDKFENFSFALIMLVFGILVNI 378

RESULT 10

TI6070
hypothetical protein F14B8.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
Accession: T16070
Geisel, C.
submitted to the EMBL Data Library, June 1995
A:Description: The sequence of C. elegans cosmid F14B8.
A:Reference number: Z18456
A:Accession: T16070
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-453 <EIT>
A:Cross-references: EMBL:028737; NID:9860717; PID:9860719; PIDN:AAA68272.1; CESP:F14B8.3
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:F14B8.3
A:introns: 56/3; 193/2; 224/3; 289/2; 302/3; 370/2; 402/3; 442/3

Query Match
Best Local Similarity 9.5%; Score 96; DB 2; Length 455;
Matches 30; Conservative 15; Mismatches 33; Indels 20; Gaps 3;

Db 49 ILCTIQLFSGIM-----FSGVITFLTKKYP--FPPIFLSGYFPG 91

Db 303 LLDLVEIYFGLISRLNHLISLCINGISMCASGVLIIFGLQKNPPIYLVFLIGY---A 359

Db 92 SYLFINSAGFLIAVKRKTETLLISRLMNLALRAI 129

Db 360 SIEFIMDACVLCVEQVPTTEVGTGACSFSLRSLGI 397

RESULT 11

T07892
probable inorganic phosphate transport protein PT1 - barrel medic
C:Species: Medicago truncatula (barrel medic)
C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 23-Mar-2001
C:Accession: T07892
R:Jiu, H., Trieu, A.T., Blaylock, L.A., Harrison, M.J.
Mol. Plant Microbe Interact. 11, 14-22, 1998
A:Title: Cloning and characterization of two phosphate transporters from Medicago tru
A:Reference number: Z16196; MWID:98086876
A:Accession: T07892
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: mRNA
A:Residues: 1-537 <LIU>
A:Cross-references: EMBL:AF000354; NID:92465526; PIDN:AAB81346.1; PID:92465527
A:Genetics:
A:Gene: PT1
C:Superfamily: probable inorganic phosphate transport protein PHO84

Query Match
Best Local Similarity 9.4%; Score 95.5; DB 2; Length 537;
Matches 34; Conservative 22; Mismatches 50; Indels 27; Gaps 4;

Db 59 IMFSPGVITFLTKKYP-----RPFILSGYFPGSVLFINSAGFLIAVKRKTTE 111

Db 377 MGFETVTFMFGLAIDYDHMSKEENRIGFVMSLTFFSNFGPNATVPAE----- 431

Db 112 TLILSRIMNL--LSALRAIAGIILTFGFI-----LDQNYICGSHONSCKA 158

Db 432 --IIPARLRSTCHGISAAKAGAVGAFGLVAAQSKDPTKMDKGYPIGIGKNSLML 489

Db 159 VTVLFGILITLM 171

Db 490 GVINFGMLCTLL 502

RESULT 12

C70379
hypothetical protein aq_917 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: C70379
V:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MWID:98196666
A:Accession: C70379
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-408 <AQF>
A:Cross-references: EMBL:AE000713; NID:92983424; PIDN:AAC07022.1; PID:92983440; GB:AE00
A:Experimental source: strain VFS
C:Genetics:
A:Gene: aq_917

Query Match
Best Local Similarity 9.4%; Score 95; DB 2; Length 408;
Matches 51; Conservative 23; Mismatches 45; Indels 84; Gaps 10;

Db 36 QSPLOK-LFARKKKILGTIOILFGIMFES---FQVITFLTKKYP--RPFILSGYFPG 90

Db 79 EDPLKATFVHIGLTGTLPLVGLXSFNKVALLGLFFLLI----- 119

Db 91 GSVLFINSAGFLIAVKRKTETLLISRLMNLALRAI-----AGIILT----- 136

Db 120 -SLIYISS---LMIYK-----LIRIKNLDPARGFPAFGSTFAGITLITFLVNL 166

Thu Oct 11 10:24:06 2001

us-09-735-712-2.rpr

Page 6

Db 344 TVSKAOTLIA----LCSTVPGYMEFTVAFIDILGRFPIQLMGFIEMTLEMFALAIYPDHR 399
QY 78 ----RPFPILSGYPFWGSYLFINSQAFILAVKRTTETLILSRIMNL--LSALRAIA 130
Db 400 HRENRIIGFLIMYSILTMFANFGPNATTFVVPAE-----IFPARLRSTCHGISAASGA 452
QY 131 GILLTFPGFI-----LDQNYIGYSHONSQKAVTVLFLGILITLM 171
Db 453 GAIVGAFGLIXAOSDSEKTDAGYPPGIGVRNSLLMLACVNFGLIVFTLL 503

Search completed: October 7, 2001, 04:40:46
Job time: 1981 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2001, 04:40:05 ; Search time 24.63 seconds

(without alignments)
276.770 Million cell updates/sec

Title: US-09-735-712-2

Sequence: 1 MDSSTAHSPVFLVFPPEITA.....ISLPFSILGCHSEDDCDEQC 199

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SWISSPROT_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	112	16.9	291	1	CD20_MOUSE
2	168	16.6	297	1	CD20_HUMAN
3	159	15.7	243	1	FCERB_RAT
4	158.5	15.6	235	1	FCERB_MOUSE
5	123.5	12.2	244	1	FCERB_HUMAN
6	92.5	9.1	458	1	TCR_STRAG
7	92	9.1	324	1	NU1M_DROVA
8	91.5	9.0	666	1	NU5M_CHOCR
9	90	8.9	312	1	NU1M_DROSU
10	87	8.6	312	1	NU1M_DROME
11	87	8.6	382	1	RFC_SHITFL
12	85.5	8.4	458	1	TCR_BACST
13	85.5	8.4	458	1	TCR_STRPN
14	85	8.4	446	1	NU4M_DROME
15	84.5	8.3	333	1	NU2M_APILI
16	84	8.3	198	1	COP_CLOPE
17	84	8.3	590	1	CANL_YEAST
18	83	8.2	291	1	YC38_PORPU
19	82	8.1	578	1	AC22_STRCO
20	82	8.1	652	1	NU5M_PODAN
21	82	8.1	669	1	NU5M_ARATH
22	82	8.1	704	1	LCRD_YEREN
23	81.5	8.0	252	1	MPUL_DROME
24	81.5	8.0	365	1	CVR_ASCSU
25	81	8.0	664	1	NU5C_STNP2
26	80	7.9	319	1	YWRB_BACSU
27	80	7.9	397	1	YLH2_CAEEL
28	79.5	7.8	241	1	TSN1_HUMAN
29	79.5	7.8	346	1	YAMY_BACAD
30	79.5	7.8	447	1	NU4M_ANOGA
31	79.5	7.8	656	1	YAN9_SCHRO
32	79.5	7.8	657	1	NU5M_EMENT
33	79.5	7.8	1386	1	Y064_MYCPN

34	79	7.8	228	1	YMC_C_EMENT
35	79	7.8	324	1	NU1M_SOUAC
36	79	7.8	532	1	YAMB_BACSU
37	79	7.8	1873	1	CCAS_RABIT
38	78.5	7.7	279	1	YS95_CAEEL
39	78.5	7.7	409	1	NU4M_ASCSU
40	78.5	7.7	956	1	MTN2_MOUSE
41	78	7.7	221	1	ATP6_RHISA
42	78	7.7	347	1	NU2M_HUMAN
43	78	7.7	545	1	NU5M_ALBCO
44	78	7.7	704	1	LCRD_YERPE
45	77.5	7.6	196	1	YMO2_PARTE

ALIGNMENTS

RESULT	ID	CD20_MOUSE	STANDARD	PRT	291 AA
AC	P19437				
DT	01-FEB-1991 (Rel. 17, Created)				
DT	01-FEB-1991 (Rel. 17, Last sequence update)				
DT	01-OCT-2000 (Rel. 40, Last annotation update)				
DE	B-CELL SURFACE PROTEIN CD20 HOMOLOG (B-CELL DIFFERENTIATION ANTIGEN LY-44).				
DE	LY-44).				
GN	MS4A2 OR CD20 OR LY-44.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89067519; PubMed=2461992;				
RA	Tedder T.F., Klejman G., Distche C.M., Adler D.A., Schloessman S.F., Saito H.;				
RT	"Cloning of a complementary DNA encoding a new mouse B lymphocyte differentiation antigen, homologous to the human B1 (CD20) antigen, and localization of the gene to chromosome 19.";				
RL	J. Immunol. 141:4388-4394(1988).				
CC	- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE REGULATION OF B-CELL ACTIVATION AND PROLIFERATION.				
CC	- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.				
CC	- PTM: PHOSPHORYLATED (BY SIMILARITY).				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).				
CC	EMBL: M62541; AAA37394.1; -				
DR	PIR: A30558; A30558.				
DR	MGI: 88321; Gd20.				
KW	B-cell; Transmembrane; Phosphorylation.				
FT	DOMAIN 1 44				
FT	TRANSMEM 45 65				
FT	TRANSMEM 69 89				
FT	TRANSMEM 112 132				
FT	TRANSMEM 183 203				
FT	DOMAIN 204 291				
SQ	SEQUENCE 291 AA: 31958 MW: Df478ECD25C16FC CRC64;				
Query Match 16.9%; Score 172; DB 1; Length 291;					
Best local Similarity 29.4%; Pred. No. 4.2e-08;					
Matches 53; Conservative 29; Mismatches 64; Indels 34; Gaps 7;					
QY	32	TESTSTPDLKLFARKKKIIIGTIIQIFGIMTFSEGVILFTLTKPYRPPPIFLS-GYPRW 90			
DB	27	TSSLVGPPTOSFFWRSEKALGAVOIMNGLPHITLIGL-----LMIPGVFAPICLISVWYPLM 82			

QY 91 GSVLFNSGAFLLAVKRTETETIIISR-IMNLSALRAIAGIIL-----LTFEGFLD 142
 DB 83 GGIMTITSSGLIAAAAKTKSRKSLVRAKAVIMSSLSLEFAAISGIIITLMDILNMLSHFLK 142
 OY 143 QN-----YIC---GYSHONS---QCKAVTVLFGIITLTMTFSIIELEFIS 181
 DB 143 MRRELIQTSKPYVDIYDCEPSNSENKSPSTGYCNSIQSVFLGIIISAMLIISAFQRLVTV 202

RESULT 2
 CD20_HUMAN STANDARD; PRT; 297 AA.
 ID CD20_HUMAN P11836; P08984; Q13963;
 AC P11836; P08984; Q13963;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE B-LYMPHOCYTE ANTIGEN CD20 (B-LYMPHOCYTE SURFACE ANTIGEN BI) (LEU-16) (BP35).
 MS4A2 OR CD20.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NC NCBL_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88258386; PubMed=3260267;
 RA Stamenkovic I., Seed B.;
 RT "Analysis of the cDNA clones encoding the B lymphocyte antigen CD20 (Bl, Bp35), a type III integral membrane protein.";
 RT J. Exp. Med. 167:1973-1980(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88124792; PubMed=2448768;
 RA Tedder T.F., Streuli M., Schlossman S.F., Saito H.;
 RT "Isolation and structure of a cDNA encoding the BI (CD20) cell-surface antigen of human B lymphocytes.";
 RT Proc. Natl. Acad. Sci. U.S.A. 85:208-212(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89176281; PubMed=2466899;
 RA Tedder T.F., Klejman G., Schlossman S.F., Saito H.;
 RT "Structure of the gene encoding the human B lymphocyte differentiation antigen CD20 (Bl).";
 RT J. Immunol. 142:2560-2568(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88283639; PubMed=2456210;
 RA Einfield D.A., Brown J.P., Valentine M.A., Clark E.A., Ledbetter J.A.;
 RT "Molecular cloning of the human B cell CD20 receptor predicts a hydrophobic protein with multiple transmembrane domains.";
 RT EMBO J. 7:711-717(1988).
 CC -1- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE REGULATION OF B-CELL ACTIVATION AND PROLIFERATION.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- PTM: PHOSPHORYLATED. MIGHT BE FUNCTIONALLY REGULATED BY PROTEIN KINASE(S).
 CC -1- DATABASE: NAME=PROT; NOTE=CD guide CD20 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd20.htm".
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC EMBL: X12530; CAA31046.1; -
 CC EMBL: M27394; AAA35581.1; -
 CC EMBL: M27395; - NOT_ANNOTATED_CDS.
 CC EMBL: L23419; AAA88911.1; -
 CC EMBL: L23415; AAA88911.1; JOINED.

DR EMBL: L23416; AAA88911.1; JOINED.
 DR EMBL: L23417; AAA88911.1; JOINED.
 DR EMBL: X07203; CAA30178.1; -
 DR EMBL: X07204; CAA30180.1; -
 DR PIR: A27400; A27400.
 DR PIR: J10042; J10042.
 DR PIR: A30586; A30586.
 DR PIR: S00387; S00387.
 DR MIM: 112210; -
 KW B-cell; Transmembrane; Phosphorylation.
 FT DOMAIN 1 63 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 64 84 POTENTIAL.
 FT TRANSMEM 85 105 POTENTIAL.
 FT TRANSMEM 121 141 POTENTIAL.
 FT TRANSMEM 189 209 POTENTIAL.
 FT DOMAIN 210 297 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 81 167 PROBABLE.
 FT DISULFID 167 183 PROBABLE.
 FT DISULFID 111 220 PROBABLE.
 FT CONFLICT 13 13 P -> L (IN REF. 4).
 FT CONFLICT 71 71 M -> I (IN REF. 3).
 SQ SEQUENCE 297 AA; 33077 MW; AC5420F8B626BDD1 CRC64;

Query Match 16.6%; Score 168; DB 1; Length 297;
 Best Local Similarity 30.3%; Pred. No. 9.6e-08;
 Matches 54; Conservative 27; Mismatches 63; Indels 34; Gaps 7;

QY 34 STQSPLOKLARKKILGTQILFGITFSFQVIFLTLKPPRRF-PFIFLGYPPWGS 92
 DB 36 SLVGPTQSFPMRESKTLGAVQINGLGFHIALGGI---LMIPGAYAVICVTWYPLMG 91
 QY 93 VLFINSGAFLLAVKRTETETIIISRIMNLSALRAIAGIIL-----T 136
 DB 92 IMYITSSLLAAETKRNKCLVGNKMTMSLSPAISGIIITLMDILNMLSHFLKMS 151
 QY 137 FGFLLDN-YICG-----SHONS---QCKAVTVLFGIITLTMTFSII-ELFIS 181
 DB 152 LNFIRAHPTPIININCEPANPSEKNSPSTGYCISGLFGLISVMLIFAFQELVIA 209

RESULT 3
 FCER_RAT STANDARD; PRT; 243 AA.
 ID FCER_RAT P13386;
 AC P13386;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR BETA-SUBUNIT (FCER1) (IGE FC RECEPTOR, BETA-SUBUNIT).
 GN MS4A2 OR FCER1B OR FCER1B.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NC NCBL_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 11-29; 37-43 AND 217-243.
 RX MEDLINE=88320465; PubMed=2970642;
 RA Kinet J.-P., Blank U., Ra C., White K., Metzger H., Kochan J.;
 RT "Isolation and characterization of cDNAs coding for the beta subunit of the high-affinity receptor for immunoglobulin E.";
 RT Proc. Natl. Acad. Sci. U.S.A. 85:6483-6487(1988).
 CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN EPSILON. HIGH AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE) RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
 CC -1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO DISULFIDE LINKED GAMMA CHAINS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration


```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC of the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/or-send-an-email-to-license@sib-sib.ch).
```

```
CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
```

```
CC EMBL; J05019; AAA37601.1; -.
DR PIR; B34342; B34342.
KW MGD; MGI:95495; Fcεr1b.
CC Ige-binding protein; Receptor; Transmembrane.
CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1 51
FT TRANSMEM 52 71
FT DOMAIN 72 89
FT TRANSMEM 90 109
FT DOMAIN 110 122
FT TRANSMEM 123 142
FT DOMAIN 143 171
FT TRANSMEM 172 191
FT DOMAIN 192 235
SQ SEQUENCE 235 AA; 25963 MW; 1C2D8BF195738168 CRC64;
```

```
Query Match 15.6%; Score 158.5; DB 1; Length 235;
Best Local Similarity 30.5%; Pred.No. 5.1e-07;
Matches 60; Conservative 29; Mismatches 81; Indels 27; Gaps 9;
```

```
OY 1 MDSSAHSPVPELVFP-PETTASEYSTELSTTFSTQSLQ-KLFARK-MKILGTIOIL 56
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 1 MDTEN-RERADLALEPNPOESSAPDIELLESPKAAPKKRTWTLTKLELPFGARDIOL 59
OY 57 FGIMTFSGVFLEFLTLKPYP---PRPFIFLSGYPPFWGSVLEFNSGAFLIAVRKTTTE 111
|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 60 VGLLICLCGTITVCISVL---YSDPEDEEVLLLYKLGGYPPMGAVLEVLSGFLSIISERNKTL 116
OY 112 TLIIISRMLNLSLARATAGIILLFFGITLDONITCGSHONSCKAKT-----VLF 164
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 117 YIVRGSIGANIVSSIAAQTGIALMIL-----NLTNPAVMNN-GKNATEDDGCFVASFT 169
OY 165 GLITLMTFSEILEFIS 181
||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 170 TELVLMMLEFLITIAFCS 186
```

```
RESULT 5
FCεB_HUMAN STANDARD: PRT: 244 AA.
```

```
ID FCεB_HUMAN ID FCεB_HUMAN
AC Q01362; DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR BETA-SUBUNIT (FCεRI)
DE (IGE FC RECEPTOR, BETA-SUBUNIT).
CN MS4A2 OR FCεRIβ.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxId=9606; [1]
RN RP SEQUENCE FROM N.A.
RX MEDLINE=92316966; PubMed=1535625;
RA Kuuster H., Zhang L., Britti A.T., Macgiaslan D.W., Kinet J.-P.;
RT "The gene and cDNA for the human high affinity immunoglobulin E
receptor beta chain and expression of the complete human receptor.";
RL J. Biol. Chem. 267.12782-12787(1992). [2]
RN RP SEQUENCE FROM N.A.
RX MEDLINE=92339505; PubMed=1386024;
RA Maekawa K., Imagawa N., Tanaka Y., Harada S.;
RT "Determination of the sequence coding for the beta subunit of the
human high-affinity Ige receptor.";
RL FEBS Lett. 302:161-165(1992). [3]
RN RP VARIANT GLU-237.
```

RA MEDLINE=96414302; PubMed=8817330.
RA Hill M.R., Cookson W.O.;
RT "A new variant of the beta subunit of the high-affinity receptor for
RT immunoglobulin E (Fc epsilon RI-beta E2376): associations with
RT measures of atopy and bronchial hyper-responsiveness.";
RL Hum. Mol. Genet. 5:959-962(1996).
RN [4]
RP VARIANT GLU-237.
RX MEDLINE=96440420; PubMed=8842731;
RX Shirakawa T., Mao X.-Q., Sasaki S., Enomoto T., Kawai M., Morimoto K.,
RA Hopkin J.;
RT "Association between atopic asthma and a coding variant of Fc-epsilon-
RT Ri-beta in a Japanese population.";
RL Hum. Mol. Genet. 5:1129-1130(1996).
RN [5]
RP ERRATUM.
RX MEDLINE=97123518; PubMed=8968765;
RX Shirakawa T., Mao X.-Q., Sasaki S., Enomoto T., Kawai M., Morimoto K.,
RA Hopkin J.;
Hum. Mol. Genet. 5:2068-2068(1996).
-1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN EPSILON. HIGH
AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC
RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL
ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)
RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR
ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
-1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO
DISULFIDE LINKED GAMMA CHAINS.
-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
-1- TISSUE SPECIFICITY: FOUND ON THE SURFACE OF MAST CELLS AND
BASOPHILS.
-1- POLYMORPHISM: VARIANT GLU-237 HAS BEEN FOUND TO BE PRESENT IN ABOUT
5.3% OF A 1004 INDIVIDUALS POPULATION SAMPLE IN AUSTRALIA. IT
SEEMS TO BE A RISK FACTOR FOR ATOPIC DERMATITIS AND ASTHMA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation --
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: D10583; BAA01440.1; -;
DR EMBL: M89796; AAA60269.1; -;
DR PIR: S21154; S21154.
DR PIR: A42806; A42806.
RN MIM: 147138; -;
RT Ige-binding protein; Receptor; Transmembrane; Polymorphism.
FT DOMAIN 1 59
FT TRANSMM 60 79
FT TRANSMM 80 97
FT DOMAIN 98 117
FT TRANSMM 118 130
FT TRANSMM 131 150
FT DOMAIN 151 180
FT TRANSMM 181 200
FT DOMAIN 201 244
FT VARIANT 237 237
FT
FT /FTID=VAR_003965.
SEQUENCE 244 AA: 26533 MW: CB523102D5F567AF CRC64;

Query Match	12.2%	Score 123.5;	DB 1;	Length 244;
Best Local Similarity	25.6%;	Pred. No. 0.00056;		
Matches 50;	Conservative 33;	Mismatches 93;	Indels 19;	Gaps.
Qy	9	PVFLVPEPEIASEYESTELSATFTSTQSPQ---	KLFARKMKILGTQIQLEFGIMTFSEFG	65
Db	21	PAFEVL--EISPOEVSSRLKLS--ASSPLHMLTVLKRQGEFLQYLITANICLCFG	76	
Oy	66	VLEFLTLKPPRPPEFT--FLSGIPFQGSVLFINSGAFLVAKRKTTETLIISRIINML	123	

Db 77 TWGVSVDIHSIEDIFFESSFRAGPEFGAIFPFISSGLMSTIISERNATVLYRGLSANTA 136

QY 124 SALAALINGIIITLTCFGLIDQYICGFSHONS----CKNAVTVLEGLITLMTF--SIIE 177

Db 137 SSINGAGCITL-----IINLKSLATYHIHSCQCFEFTKCFMASFSTREIYVMMLFLITLG 192

QY 178 LFISLPSIILGCHSE 192

Db 193 LGSAVSLTIGAGEE 207

RESULT	6
TCCR_STRAG	
ID	TCCR_STRAG
AC	P13924;
DT	01-JAN-1990 (Rel. 13, Created)
DT	01-JAN-1990 (Rel. 13, Last sequence update)
DT	01-OCT-1994 (Rel. 30, last annotation update)
DE	TETRACYCLINE RESISTANCE PROTEIN.
GN	TEM.
OS	Streptococcus agalactiae.
OG	Plasmid pMV158.
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
CC	Streptococcus.
XX	NCBI_TaxID=1311;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=90016790; PubMed=2677995;
RA	van der Lelle D., Bron S., Venema G., Oskam L.;
RT	"Similarity of minus origins of replication and flanking open reading frames of plasmids pub110, pPB913 and pMV158." ;
RL	Nucleic Acids Res. 17:7283-7294(1989)
CC	-1- FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN ACCUMULATES AS A METAL-TETRACYCLINE/H+ ANTIPORTER.
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC	-1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isdb-sib.ch/announcement/ or send an email to license@isdb-sib.ch).
CC	-----
DR	EMBL; X15669; CAA3712.1; -
DR	PIR; G25599; YTSOG.
DR	InterPro: IPR001411; -
DR	PRINTS; PR01036; TCRTETB.
KW	Antibiotic resistance; Transmembrane; Transport; Symport; Plasmid.
FT	TRANSSEM 12 33 POTENTIAL.
FT	TRANSSEM 81 100 POTENTIAL.
FT	TRANSSEM 111 129 POTENTIAL.
FT	TRANSSEM 140 162 POTENTIAL.
FT	TRANSSEM 165 185 POTENTIAL.
FT	TRANSSEM 201 221 POTENTIAL.
FT	TRANSSEM 223 240 POTENTIAL.
FT	TRANSSEM 256 276 POTENTIAL.
FT	TRANSSEM 297 317 POTENTIAL.
FT	TRANSSEM 324 344 POTENTIAL.
FT	TRANSSEM 346 365 POTENTIAL.
FT	TRANSSEM 432 451 POTENTIAL.
QO	SEQUENCE 458 AA; 50006 MW; AD2014E7CA19995 CRC64;

Query Match	9.18;	Score 92.5;	DB 1;	Length 456;
Best Local Similarity	25.38;	Pred. No. 0.49;		
Matches 44;	Conservative 27;	Mismatches 68;	Indels 35;	Gaps 7;
OY	32	TESTQSPLOKTFARKMILGTLQILFGIMTFSEGVILFTLLKPYPPRPFLFSGYPMWG	91	

```

Db 178 TITVPEMLMLKKEVRKIKHFDIK-GILMSVGIVF-----MLFTSYSTI-- 223
QY 92 SVLFINSGAFLIAVK--RKTEFLILISRLMMLLSALRAIAIGILL--TFGFLIDQNYIC 147
Db 224 SFLIVSVLSLSEFLIEFKKIKRTYDFVDPGLCKNILEMIGVLCGGIFETVAGFVSMVPM 283
QY 148 GYSHONSQCKAVNYLF-----GILL-----TMTFSIIEFLFSLRP 184
Db 284 KDVAHQLSTAEIGSVIIFPGTMSVIIIEGIGILVDRGRGPLYVNIIGVTLFVSVP 337

RESULT 7
NUM_DROVA STANDARD: PRT; 324 AA.
AC P07710;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 1 (EC 1.6.5.3).
MT:ND1 OR ND1.
Drosophila yakuba (Fruit fly).
Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7245;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=2317.6 IVORY COAST;
RC MEDLINE=86089137; PubMed=3001325;
RA Clary D.O., Wolstenholme D.R.;
RT "The mitochondrial DNA molecular of Drosophila yakuba: nucleotide
RT sequence, gene organization and genetic code."
RT J. Mol. Evol. 22:252-271(1985).
RN [2];
RP SEQUENCE FROM N.A.
RC MEDLINE=84221393; PubMed=6328435;
RA Clary D.O., Wolstenholme D.R.;
RT "Sequence and arrangement of the genes for cytochrome b, URF1, URF4,
RT URF4, URF5, URF6 and five tRNAs in Drosophila mitochondrial DNA."
RT Nucleic Acids Res. 12:3747-3762(1984).
CC 1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
CC 1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X03240; CA26997.1; -
DR PIR: D30020; D30020.
DR FLYBase; FBgn0013183; Dyak\mt:ND1.
DR InterPro: IPR001694; -
DR Pfam: PF00146; NADhdh; 1.
DR PROSITE: PS00667; COMPLEX1_ND1_1; 1.
DR PROSITE: PS00668; COMPLEX1_ND1_2; 1.
DR Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
SQ SEQUENCE 324 AA; 37529 MW; D3F20A5968463196 CRC64;

Query Match 9.1%; Score 92; DB 1; Length 324;
Best local Similarity 26.3%; Pred. No. 0.39;
Matches 50; Conservative 23; Mismatches 63; Indels 54; Gaps 9;

QY 48 KILGTIOI-----LFGIMFSEFVIFLTKPKYPRP-----FIFISGY- 87
Db 30 KVLGVIQIRKPKVKGIMGPFCALIKLFTKEQITPLISNLYSYISFISFLSLV 89
QY 88 ----PFWGSVLFIN-SGAFLIAVKRKTTETLLISRLIM-----LISALRAIA----- 130

```

```

Db 90 WMCPFFVYKLSFNLGGLFLCCSLGVTVMAGMSNSNYALGGLRAVAQRTSYEVS 149
QY 131 -GIILLTFEFTL-DQNYICGYSHONSQCKAVNYLFILITLMTFSIIEFLISF----- 182
Db 150 LALIMLSFILLGSGYNNMIVFFYQ-----IYMFLLIFPMISLWMLTISLAETNRT 200
QY 183 PFTSLGCHSE 192
Db 201 PFDPAEGESE 210

RESULT 8
NUM_CHOCR STANDARD: PRT; 666 AA.
AC P48920;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 5 (EC 1.6.5.3).
GN ND5 OR NAD5.
OS Chondrus crispus (Carraheen).
OG Mitochondrion.
OC Eukaryota; Rhodophyta; Florideophyceae; Gigartinales; Gigartinales;
OC Chondrus.
OX NCBI_Taxid=2769;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=Apices;
RC MEDLINE=95341681; PubMed=7616569;
RA Leblanc C., Boyen C., Richard O., Bonnard G., Grienenberger J.M.,
RA Kloeareg B.;
RT "Complete sequence of the mitochondrial DNA of the rhodophyte
RT Chondrus crispus (Gigartinales). Gene content and genome
RT organization."
RT J. Mol. Biol. 250:484-495(1995).
CC 1- CATALYTIC ACTIVITY: NADH + ACCEPTOR = NAD(+) + REDUCED ACCEPTOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z47547; CA87625.1; -
DR Mendel; 7732; CHOCR:nd5;1.
DR InterPro: IPR001516; -
DR InterPro: IPR001750; -
DR Pfam: PF00361; oxidored_q1; 1.
DR Pfam: PF00662; oxidored_q1_N; 1.
DR Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 666 AA; 75628 MW; 37E86F2C24B9D360 CRC64;

Query Match 9.0%; Score 91.5; DB 1; Length 666;
Best local Similarity 21.2%; Pred. No. 0.86;
Matches 56; Conservative 32; Mismatches 83; Indels 93; Gaps 9;

QY 16 PELTASEYESTELSAVTFSTOSPLQKLFARKMKILGTIOILFGIMFSEFVIFLTKPK 75
Db 394 PLTGFYSKDFLELQOISSYSLNQMYSISFACWLGIMSVFF-TSYYSFLIYTLFLNNT 452
QY 76 -----YRPEPFIFL-----SGYPPWGSVLFIN----- 98
Db 453 NLAQSSLNIVHSSSLMIFPLILISIGSIFAGYLIRDLFEVSGSDPFGAIFILPKHSTP 512
QY 99 -----GAFLIAVKRKTTETLLISRLMMLLSALRAIA----- 131
Db 513 IEAEELPIYWKMLPFLISLIGLFFASFPVQIFLTKTFYFKSNLQNLSEFFLLIKKKWYMDV 572
QY 132 -----ILLTFPGF-----ILDONRY-----CGYSHONS-----QCKAVTVLFLG 165

```



```

RFC_SHIFL
ID RFC_SHIFL STANDARD: PRT: 382 AA.
AC P37784;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE O-ANTIGEN POLYMERASE.
GN RFC.
OS Shigella flexneri.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SEROTYPE 2A;
RX MEDLINE=94131953; PubMed=7507920;
RA Morona R., Mavris M., Fallarino A., Manning P.A.;
RT "Characterization of the rfc region of Shigella flexneri.";
RL J. Bacteriol. 176:733-747(1994).
CC -1- FUNCTION: MAY LINK THE O-ANTIGEN TETRASACCHARIDE UNITS INTO LONG
CHAINS, GIVING RISE TO TYPICAL SMOOTH LPS.
CC -1- PATHWAY: LIPOLYSACCHARIDE BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
(POTENTIAL).
CC -1- SIMILARITY: TO OTHER O-ANTIGEN POLYMERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X71970; CAA50774.1;
KW Lipopolysaccharide biosynthesis; Transmembrane; Inner membrane.
FT TRANSMEM 11 31 POTENTIAL.
FT TRANSMEM 36 56 POTENTIAL.
FT TRANSMEM 61 81 POTENTIAL.
FT TRANSMEM 87 107 POTENTIAL.
FT TRANSMEM 126 146 POTENTIAL.
FT TRANSMEM 166 186 POTENTIAL.
FT TRANSMEM 194 214 POTENTIAL.
FT TRANSMEM 228 248 POTENTIAL.
FT TRANSMEM 303 323 POTENTIAL.
FT TRANSMEM 325 345 POTENTIAL.
FT TRANSMEM 348 368 POTENTIAL.
SQ SEQUENCE 382 AA; 43723 MW; DCE02BB41619156A CRC64;

Query Match 8.6%; Score 87; DB 1; Length 382;
Best Local Similarity 19.7%; Pred. No. 1.2;
Matches 46; Conservative 39; Mismatches 64; Indels 84; Gaps 8;

```

```

ID TCR_BACST STANDARD: PRT: 458 AA.
AC P07561;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE TETRACYCLINE RESISTANCE PROTEIN.
GN TET.
OS Bacillus stearohermophilus.
OC plasmid pTHR15.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86031344; PubMed=2996983;
RA Hoshino T., Ikeda T., Tomizuka N., Furukawa K.;
RT "Nucleotide sequence of the tetracycline resistance gene of pTHR15, a
RT thermophilic Bacillus plasmid: comparison with staphylococcal tcr
RT controls.";
RL Gene 37:131-138(1985).
CC -1- FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE
EFFECT. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE
ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN
CC FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTI-PORTER.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M1036; AAA22851.1;
DR PIR: A23973; YTBSTR.
DR InterPro: IPR001411;
DR PRINTS: PRO1036; TCRTEB.
KW Antibiotic resistance; Transmembrane; Transport; Symport; Plasmid.
FT TRANSMEM 12 33 POTENTIAL.
FT TRANSMEM 81 100 POTENTIAL.
FT TRANSMEM 111 129 POTENTIAL.
FT TRANSMEM 140 162 POTENTIAL.
FT TRANSMEM 165 185 POTENTIAL.
FT TRANSMEM 201 221 POTENTIAL.
FT TRANSMEM 223 240 POTENTIAL.
FT TRANSMEM 256 276 POTENTIAL.
FT TRANSMEM 297 317 POTENTIAL.
FT TRANSMEM 324 344 POTENTIAL.
FT TRANSMEM 346 365 POTENTIAL.
FT TRANSMEM 432 451 POTENTIAL.
SQ SEQUENCE 458 AA; 50119 MW; E1F3F4C57E7B06A CRC64;

Query Match 8.4%; Score 85.5; DB 1; Length 458;
Best Local Similarity 24.7%; Pred. No. 2;
Matches 43; Conservative 27; Mismatches 69; Indels 35; Gaps 7;

```

RESULT 12
TCR_BACST

```

RESULT 13
TCR_STRPN STANDARD: PRT: 458 AA.
ID TCR_STRPN STANDARD: PRT: 458 AA.
AC P11063; P72219;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE TETRACYCLINE RESISTANCE PROTEIN.
GN TET.
OS Streptococcus pneumoniae, Bacillus cereus, and Bacillus subtilis.
OG Plasmid pSL1, Plasmid pBC16, plasmid pHY163PLK, Plasmid pTB19, and
OG Plasmid pMS1981.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus
OX NCBI_TaxID=1313, 1396, 1423;
RN [1]
RN SEQUENCE FROM N.A.
RP SPECIES=S.pneumoniae; PLASMID=PLS1;
RC MEDLINE=8726167; PubMed=2438417;
RC Lacks S.A., Lopez P., Greenberg B., Espinosa M.;
RC "Identification and analysis of genes for tetracycline resistance and
RC replication functions in the broad-host-range plasmid pSL1.";
RC J. Mol. Biol. 192:753-765(1986).
RN [2]
RN SEQUENCE FROM N.A.
RP SPECIES=B.cereus; PLASMID=PBC16;
RC MEDLINE=90221899; PubMed=2109312;
RC Palva A., Vildgren G., Simonen M., Rintala H., Laamanen P.;
RC "Nucleotide sequence of the tetracycline resistance gene of pBC16
RC from Bacillus cereus.";
RC Nucleic Acids Res. 18:1635-1635(1990).
RN [3]
RN SEQUENCE FROM N.A.
RP PLASMID=PHY163PLK;
RC Ishiwa H., Shubahara H.;
RC "New shuttle vectors for Escherichia coli and Bacillus subtilis. III.
RC Nucleotide sequence analysis of tetracycline resistance gene of
RC pAMalpha1 and ori-177.";
RC Jpn. J. Genet. 60:485-498(1985).
RN [4]
RN SEQUENCE FROM N.A.
RP PLASMID=PHY163PLK;
RC MEDLINE=83129391; PubMed=6186390;
RC MEDLINE=92052681; PubMed=1946749;
RC Oskam L., Hillenga D.J., Venema G., Bron S.;
RC "The large Bacillus plasmid pTB19 contains two integrated
RC rolling-circle plasmids carrying mobilization functions.";
RC Plasmid 26:30-39(1991).
RN [5]
RN SEQUENCE FROM N.A.
RP SPECIES=B.subtilis; PLASMID=PTB19;
RC MEDLINE=92052681; PubMed=1946749;
RC Oskam L., Hillenga D.J., Venema G., Bron S.;
RC "The large Bacillus plasmid pTB19 contains two integrated
RC rolling-circle plasmids carrying mobilization functions.";
RC Plasmid 26:30-39(1991).
RN [6]
RN SEQUENCE FROM N.A.
RP SPECIES=B.subtilis; STRAIN=168 / MABBURG; PLASMID=pMS1981;
RC MEDLINE=86287699; PubMed=3090576;
RC Saeguchi R., Shishido K., Hoshino T., Furukawa K.;
RC "The nucleotide sequence of the tetracycline resistance gene of
RC plasmid pMS1981 from Bacillus subtilis differs from pTB19 from a
RC thermophilic Bacillus by two base pairs.";
RC Plasmid 16:72-73(1986).
RN [7]
RN PUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE
RN EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE
RN ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN
RN FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTI-PORTER.
RN -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
RN -1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
RN AS THE DRUG RESISTANCE TRANSCOASE FAMILY).
RN
RN THIS SWISS-PROT entry is copyright It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X51366; CAA35751.1; -
DR EMBL; M29725; AAA98167.1; -
DR EMBL; D13792; -; NOT_ANNOTATED_CDS.
DR EMBL; M63891; AAA98304.1; -
DR EMBL; D000006; BAA00005.1; -
DR PIR; S09234; YTSU6.
DR PIR; C25599; YTSOG.
DR PIR; J01211; J01211.
DR InterPro; IPR001411. -
DR PRINTS; PR01036; TCTETTB.
KW Antibiotic resistance; Transmembrane; Transport; Symport; Plasmid.
FT TRANSMEM 12 33 POTENTIAL.
FT TRANSMEM 81 100 POTENTIAL.
FT TRANSMEM 111 129 POTENTIAL.
FT TRANSMEM 140 162 POTENTIAL.
FT TRANSMEM 165 185 POTENTIAL.
FT TRANSMEM 201 221 POTENTIAL.
FT TRANSMEM 223 240 POTENTIAL.
FT TRANSMEM 256 276 POTENTIAL.
FT TRANSMEM 297 317 POTENTIAL.
FT TRANSMEM 324 344 POTENTIAL.
FT TRANSMEM 346 365 POTENTIAL.
FT TRANSMEM 432 451 POTENTIAL.
SO SEQUENCE 458 AA; 50092 MW; 6A70B777D44C2074 CRC64;

```

Query Match	8.4%	Score 85.5	DB 1	Length 458
Best Local Similarity	24.7%	Pred. No. 2		
Matches 43	Conservative 27	Mismatches 69	Indels 35	Gaps 7

QY	32	TESTOSLOKLFARKMKLGTIOLFGIMTFSGFVIEFLTLKPYRPRPFILSGYPMWG	91
DB <td>178</td> <td>TIITPYFLMKLTKKEVRKIGHFDIK-GIILMSGVGF-----MLFTSYST--</td> <td>223</td>	178	TIITPYFLMKLTKKEVRKIGHFDIK-GIILMSGVGF-----MLFTSYST--	223
QY <td>92</td> <td>SVLFINSGAPLIAVK--RKTEYLLISRMNLLSALRAGIILL--TFGFILDONYIC</td> <td>147</td>	92	SVLFINSGAPLIAVK--RKTEYLLISRMNLLSALRAGIILL--TFGFILDONYIC	147
DB <td>224</td> <td>SFLIYSVLSPFLIEYKHRIKVTDPFVDPDLGKNIPDMICVLGGLIFGVAVGFVSNVPM</td> <td>283</td>	224	SFLIYSVLSPFLIEYKHRIKVTDPFVDPDLGKNIPDMICVLGGLIFGVAVGFVSNVPM	283
QY <td>148</td> <td>GYSHNSOCKAKVYVLEL-----GILL---TLMTFSIIEFLISLPF</td> <td>184</td>	148	GYSHNSOCKAKVYVLEL-----GILL---TLMTFSIIEFLISLPF	184
DB <td>284</td> <td>KDVHQLSTAEIGSVIIFPGTMSVLIIFGIGGILVDRGRLVNLGVFLSVSF</td> <td>337</td>	284	KDVHQLSTAEIGSVIIFPGTMSVLIIFGIGGILVDRGRLVNLGVFLSVSF	337

RESULT 14	NU4M_DROME	STANDARD	PRT	446 AA.
AC	P18931			
DT	01-NOV-1990 (Rel. 16, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	MDM8-BIGUINONE OXIDOREDUCTASE CHAIN 4 (EC 1.6.5.3).			
GN	MT:ND4 OR ND4.			
OG	Drosophila melanogaster (Fruit fly).			
OC	Mitochondrion.			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxId=7227;			
OX	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-BRETAGNE;			
RC	MEDLINE=882121147; PubMed=3130291;			
RA	Garesse R;			
RT	"Drosophila melanogaster mitochondrial DNA: gene organization and			
RT	evolutionary considerations.";			
TL	Genetics 118:649-663(1988).			

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2001, 04:40:46 ; Search time 37.35 Seconds
(without alignments)
303.882 Million cell updates/sec

Title: US-09-735-712-8

Perfect score: 771

Sequence: 1 MDSSTAHSPVFLVPEPPEITL.....SLPFSILGCHSDCCDCQCC 149

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

PIR-68: **
1: pirl: **
2: pirl: **
3: pirl: **
4: pirl: **

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	150.5	19.5	214	2	I59258
2	124	16.1	291	2	A30558
3	123.5	16.0	243	2	A31331
4	117.5	15.2	235	2	B34342
5	110.5	14.3	297	1	A30586
6	100	13.0	457	2	T37205
7	99.5	12.9	244	2	A42806
8	92	11.9	455	2	T16070
9	88.5	11.5	755	2	T02553
10	88	11.4	408	2	C70379
11	87	11.3	464	2	T03780
12	86.5	11.2	1431	2	T22748
13	85.5	11.1	712	2	T02552
14	84	10.9	991	2	B71315
15	83.5	10.8	555	2	F72355
16	82.5	10.7	780	2	H84685
17	82	10.6	657	2	S04724
18	81.5	10.6	590	1	ORBYPR
19	81	10.5	228	2	F71886
20	80.5	10.4	266	2	T41414
21	80.5	10.4	662	2	S62707
22	80	10.4	458	1	YTBSTR
23	80	10.4	458	1	YTBSTR
24	80	10.4	458	1	YTBSTR
25	80	10.4	458	2	J01211
26	79.5	10.3	244	2	H70193
27	79	10.2	292	2	C70421
28	78.5	10.2	460	2	T13881
29	78.5	10.2	573	2	S60912

30	78.5	10.2	843	2	T06068	probable proton pu
31	77.5	10.1	144	2	C69548	conserved hypotet
32	77.5	10.1	262	2	S23241	hypothetical prote
33	77	10.0	669	2	T08827	hypothetical prote
34	76.5	9.9	482	2	T17022	NADH dehydrogenase
35	76.5	9.9	633	2	T21779	hypothetical prote
36	76	9.9	152	2	JH0751	IgE receptor beta
37	76	9.9	475	2	T26850	hypothetical prote
38	75.5	9.8	307	2	C86844	hypothetical prote
39	75.5	9.8	381	2	F59101	hypothetical prote
40	75.5	9.8	669	1	DNM005	NADH dehydrogenase
41	75.5	9.8	768	2	S52684	probable membrane
42	75.5	9.8	803	1	E70041	probable copper-tr
43	75	9.7	306	2	H83620	probable permease
44	75	9.7	365	2	F82210	amino acid ABC tra
45	75	9.7	601	2	T13820	NADH dehydrogenase

ALIGNMENTS

RESULT 1
I59258
IgE receptor beta chain / CD20 antigen homolog - human
C:Species: Homo sapiens (man)
C:Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 05-Nov-1999
C:Accession: I59258
R:Adra, C.N.; Leilas, J.M.; Kobayashi, H.; Kaghad, M.; Morrison, P.; Rowley, J.D.; Li
Proc. Natl. Acad. Sci. U.S.A. 91, 10178-10182, 1994
A:Title: Cloning of the cDNA for a hematopoietic cell-specific protein related to CD2
nning regions.
A:Reference number: I59258; MUID:95024008
A:Accession: I59258
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-214 <RES>
A:Cross-references: GB:I35848; NID:9561638; PIDN:AAA62319.1; PID:9561639
C:Genetics:
A:Gene: GDB:CD20L; HTM4
A:Cross-references: GDB:392702
A:Map position: 11q12-11q13.1
C:Keywords: immunoglobulin receptor

Query Match 19.5% Score 150.5; DB 2; Length 214;
Best Local Similarity 31.1% Pred. No. 6.3e-08;
Matches 51; Conservative 22; Mismatches 64; Indels 27; Gaps 8;
OY 1 MDSSTAH-SPVFLVPEPPEITASEYESTELSATPSTQSPLOKLPARKKINGTIOILFGI 59
Db 11 LGSASAHGTGSESEPEELMTSYH-----PINSBPYOK---AKLOYGAIQIIMAA 60
OY 60 MTFSEGVIFLETLIKPY----PREPFIETLSGYPFWGSLFINSAGFLIAVKKRTET-- 112
Db 61 MILALGV-FGLSIQYPIYFQKHFFEFYFGYPIMGAVFPCSSGTLISVAGIKPTFMWQ 119
OY 113 -LGILTLMTFSII-ELFISLPR-----SLGCHSDCCDCQCC 148
Db 120 NSFCGMNIAATIALVGTAFILSNIAVNIQSILRSSSESPPDIC 163

RESULT 2
A30558
B-cell surface antigen CD20 homolog - mouse
N:Alternate names: B-cell differentiation antigen Ly-44
C:Species: Mus musculus (house mouse)
C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 07-Feb-1997
C:Accession: A30558
R:Redder, T.F.; Klejman, G.; Distcheche, C.M.; Adler, D.A.; Schlossman, S.F.; Saito, H.
J. Immunol. 141, 4368-4394, 1988
A:Title: Cloning of a complementary DNA encoding a new mouse B lymphocyte differentia
A:Reference number: A30558; MUID:89067519
A:Accession: A30558

Nature 402, 761-768, 1999
 A>Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*
 A:Reference number: AB4420; MUID:20083487
 A:Accession: D84734
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-755 <STO>
 A:Cross-references: GB:AE002093; NID:g3298542; PIND:AA025936.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: AT2G32540; T26B15.10
 A:Map position: 2
 A:Introns: 88/2; 194/3; 236/3; 277/3; 344/3; 390/1; 450/3; 567/3

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2001, 04:46:09 ; Search time 24.63 Seconds

(Without alignments)
207.230 Million cell updates/sec

Title: US-09-735-712-8

Perfect score: 771
1 MDSSTAHSPVLPVFPPEITA.....SLPESILCHSEDCDEQCC 149

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	124	16.1	291	1	CD20_MOUSE
2	123.5	16.0	243	1	FCBE_RAT
3	117.5	15.2	235	1	FCBE_MOUSE
4	110.5	14.3	297	1	CD20_HUMAN
5	99.5	12.9	244	1	FCBE_HUMAN
6	82	10.6	657	1	NU5M_EMENT
7	81.5	10.6	590	1	CAN1_YEAST
8	80	10.4	458	1	TCR_BACST
9	80	10.4	458	1	TCR_STRAG
10	80	10.4	458	1	TCR_STRPN
11	78.5	10.2	573	1	ALP1_YEAST
12	77.5	10.1	497	1	ANSP_YEAST
13	77.5	10.0	669	1	COTE_HUMAN
14	75.5	9.8	669	1	NU5M_ARATH
15	75.5	9.8	803	1	ATCU_BACSU
16	74.5	9.7	641	1	NU5M_ALIMA
17	74.5	9.7	652	1	NU5M_PODAN
18	74	9.6	402	1	YCB_BACSU
19	73.5	9.5	605	1	NU5M_CHICK
20	73.5	9.5	522	1	YOW3_CAEEL
21	73	9.4	439	1	LNT_AOUAE
22	72.5	9.4	438	1	CKR4_HUMAN
23	72.5	9.4	438	1	MAE1_SCHPO
24	72	9.3	458	1	TCR_STRAY
25	72	9.3	499	1	ANSP_ECOLI
26	72	9.3	664	1	NU5M_PHYIN
27	71.5	9.3	356	1	EMRD_ECOLI
28	71.5	9.3	459	1	NU4M_MOUSE
29	71	9.2	315	1	NU1M_ARTSF
30	71	9.2	315	1	PSTC_HAEIN
31	71	9.2	373	1	CKR2_RAT
32	71	9.2	562	1	FLA1_METUA
33	71	9.2	606	1	NU5M_EQUAS

34	71	9.2	854	1	VP2_BOVIN	097681 bos taurus
35	70.5	9.1	166	1	YC4B_METUA	P81230 methanococ
36	70.5	9.1	349	1	FM2_MACMU	P79191 macaca mula
37	70.5	9.1	460	1	NU4M_PIG	079881 sus scrofa
38	70.5	9.1	527	1	NU2M_ACACA	037376 acanthamoeb
39	70	9.1	436	1	SECY_METUA	060175 methanococ
40	70	9.1	604	1	NU5M_HORSE	P48566 equus cabal
41	70	9.1	606	1	NU5M_SHEEP	078756 ovis aries
42	69.5	9.0	121	1	YG18_YEAST	P53151 saccharomyc
43	69	8.9	175	1	Y433_METUA	057875 methanococ
44	69	8.9	458	1	TCR_BACSU	P23054 bacillus su
45	68.5	8.9	334	1	SRB7_CAEEL	P54142 caenorhabdi

ALIGNMENTS

RESULT	ID	CD20_MOUSE	STANDARD	PRT	291 AA
AC	P19437				
DT	01-FEB-1991 (Rel. 17, Created)				
DT	01-FEB-1991 (Rel. 17, Last sequence update)				
DT	01-OCT-2000 (Rel. 40, Last annotation update)				
DE	B-CELL SURFACE PROTEIN CD20 HOMOLOG (B-CELL DIFFERENTIATION ANTIGEN LY-44)				
GN	MS4A2 OR CD20 OR LY-44				
OS	Mus musculus (Mouse)				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
OX	NCBI_TaxID=10090;				
RP	SEQUENCE FROM N.A.				
RP	MEDLINE=89067519; PubMed=2461992;				
RA	Tedder T.F., Kiehlman G., Distche C.M., Adler D.A., Schlossman S.F.,				
RA	Saito H.;				
RT	"Cloning of a complementary DNA encoding a new mouse B lymphocyte				
RT	differentiation antigen, homologous to the human B1 (CD20) antigen,				
RT	and localization of the gene to chromosome 19."				
RL	J. Immunol. 141:4388-4394(1988).				
CC	-1- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE REGULATION OF				
CC	B-CELL ACTIVATION AND PROLIFERATION.				
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.				
CC	-1- PTM: PHOSPHORYLATED (BY SIMILARITY).				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement. (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL: M62541; AAA37394.1; -				
DR	PIR: A30558; A30558.				
DR	MGI: 88321; G320.				
KM	B-cell; Transmembrane; Phosphorylation.				
FT	DOMAIN 1 44				
FT	TRANSMEM 45 65				
FT	TRANSMEM 69 89				
FT	TRANSMEM 112 132				
FT	TRANSMEM 183 203				
FT	DOMAIN 204 291				
SQ	SEQUENCE 291 AA; 31958 MW; DF478ECD2C5C1CFC CRC64;				

Query Match 16.1%; Score 124; DB 1; Length 291;

Best local Similarity 33.0%; Pred. No. 0.00017;

Matches 35; Conservative 18; Mismatches 47; Indels 6; Gaps 3;

QY 32 TSTSTSPLOKLFARKKKILGTLQIFGIMTSGVIFLTLKPYRFPFIFLS-GYFPW 90
DB 27 TSTLVPTGPTGFFMRKSKALGAVQIMNGLFRITLGLT---LMIPTGVAFICLSVYPLW 82

QY 91 GSVLFNSGAFLLAVKRRKTTETLGLITIMTFISLEFISLPSIL 136
 ID FCBF_MOUSE STANDARD; PRT; 243 AA.
 DB 83 GGIWYIISGSLAAAEKTSRK-SLVKAKYIMSSLSLFAISGILL 127

RESULT 2
 FCBF_MOUSE STANDARD; PRT; 243 AA.

AC P13386;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR BETA-SUBUNIT (FCER1)
 DE (IGE FC RECEPTOR, BETA-SUBUNIT).
 GN MS4A2 OR FCER1B OR FCB1B.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 11-29; 37-43 AND 217-243.
 RX MEDLINE=88320465; PubMed=2970642;
 RA Kinet J.-P., Blank U., Ra C., White K., Metzger H., Kochan J.;
 RT "Isolation and characterization of cDNAs coding for the beta subunit
 of the high-affinity receptor for immunoglobulin E.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:6483-6487(1988).
 CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULINS EPSILON. HIGH
 CC AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC
 CC RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL
 CC ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)
 CC RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR
 CC ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
 CC -1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO
 CC DISULFIDE LINKED GAMMA CHAINS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M22923; AAA41149.1; -
 DR PIR: A31231; A31231.
 DR IGE-binding protein; Receptor; Transmembrane.
 FT DOMAIN 1 59 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 60 79 POTENTIAL.
 FT DOMAIN 80 97 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 98 117 POTENTIAL.
 FT DOMAIN 118 130 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 131 150 POTENTIAL.
 FT DOMAIN 151 179 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 180 199 POTENTIAL.
 FT DOMAIN 200 243 CYTOPLASMIC (POTENTIAL).
 FT SEQUENCE 243 AA; 26730 MW; 47IDFA59B68B5D CRC64;

Query Match 16.0%; Score 123.5; DB 1; Length 243;
 Best Local Similarity 29.0%; Pred. No. 0.00016;
 Matches 31; Conservative 16; Mismatches 41; Indels 19; Gaps 2;

QY 36 QSPLOKFAKMKILGIIQLFGIMTFSGVIFLFTLKP--YPRPFIISGYPFGSV 93
 ID FCBF_MOUSE STANDARD; PRT; 243 AA.
 DB 47 QQWQSGSLAKLEFLGTVLVGLICLFCSTVCSITQTSDFDEVLILKRAQIPFGAV 106

QY 94 LFNISGAFLLAVKRRKT-----ETGLITIMTFIS 123
 ID FCBF_MOUSE STANDARD; PRT; 243 AA.
 DB 107 LFVLSGFLSINSRKNTLYVRSIGANIVSIAAGIAIILINLS 153

RESULT 3
 FCBF_MOUSE STANDARD; PRT; 235 AA.
 AC P20490;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR BETA-SUBUNIT (FCER1)
 DE (IGE FC RECEPTOR, BETA-SUBUNIT).
 GN MS4A2 OR FCER1B OR FCB1B.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=89359361; PubMed=2527850;
 RA C., Jouvin M.H.E., Kinet J.-P.;
 RT "Complete structure of the mouse mast cell receptor for IGE (Fc
 RT epsilon R1) and surface expression of chimeric receptors (rat-mouse-
 RT human) on transfected cells.";
 RL J. Biol. Chem. 264:15323-15327(1989).
 CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULINS EPSILON. HIGH
 CC AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC
 CC RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL
 CC ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)
 CC RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR
 CC ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
 CC -1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO
 CC DISULFIDE LINKED GAMMA CHAINS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: J05019; AAA37601.1; -
 DR PIR: B34342; B34342.
 DR MGD: MG1:93495; FcεR1b.
 DR IGE-binding protein; Receptor; Transmembrane.
 FT DOMAIN 1 51 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 52 71 POTENTIAL.
 FT DOMAIN 72 89 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 90 109 POTENTIAL.
 FT DOMAIN 110 122 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 123 142 POTENTIAL.
 FT DOMAIN 143 171 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 172 191 POTENTIAL.
 FT DOMAIN 192 235 CYTOPLASMIC (POTENTIAL).
 FT SEQUENCE 235 AA; 25963 MW; 1C2D6BF195738168 CRC64;

Query Match 15.2%; Score 117.5; DB 1; Length 235;
 Best Local Similarity 32.8%; Pred. No. 0.00054;
 Matches 39; Conservative 17; Mismatches 50; Indels 13; Gaps 6;

QY 1 MDSSTASHPELVLP--PEITASEYESTELSATTFSTQSPQ--KLPAK-WKILGIIQL 56
 ID FCBF_MOUSE STANDARD; PRT; 235 AA.
 DB 1 MDEN-RSRADLALPNQESSAPDIELLEASPAKAPQWRFTFLKLEFLGATQIL 59

QY 57 FGIMTFSGVIFLFTLKP-----PRPFIISGYPFGSVIFNSGAFLLAVKRRKT 110
 ID FCBF_MOUSE STANDARD; PRT; 235 AA.
 DB 60 VGLICLFCSTVCSITQTSDFDEVLILKRAQIPFGAV 115

RESULT 4
 ID CD20_HUMAN STANDARD; PRT; 297 AA.
 AC P11836; P08984; Q13963;

DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE B-LYMPHOCYTE ANTIGEN CD20 (B-LYMPHOCYTE SURFACE ANTIGEN B1) (LEU-16) (BP35).
 GN MS4A2 OR CD20.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88258386; PubMed=3260267;
 RA Stamenkovic I., Seed B.;
 RT "Analysis of two cDNA clones encoding the B lymphocyte antigen CD20 (B1, BP35), a type III integral membrane protein.";
 RL J. Exp. Med. 167:1975-1980(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88124792; PubMed=2448768;
 RA Tedder T.F., Streuli M., Schlossman S.F., Saito H.;
 RT "Isolation and structure of a cDNA encoding the B1 (CD20) cell-surface antigen of human B lymphocytes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:208-212(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89176281; PubMed=2466899;
 RA Tedder T.F., Klejman G., Schlossman S.F., Saito H.;
 RT "Structure of the gene encoding the human B lymphocyte differentiation antigen CD20 (B1).";
 RL J. Immunol. 142:2560-2568(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88283639; PubMed=2456210;
 RA Einfield D.A., Brown J.P., Valentine M.A., Clark E.A., Ledbetter J.A.;
 RT "Molecular cloning of the human B cell CD20 receptor predicts a hydrophobic protein with multiple transmembrane domains.";
 RL EMBO J. 7:711-717(1988).
 CC -1- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE REGULATION OF B-CELL ACTIVATION AND PROLIFERATION.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- PPM: PHOSPHORYLATED. MIGHT BE FUNCTIONALLY REGULATED BY PROTEIN KINASE(S).
 CC -1- DATABASE: NAME=PRO; NOTE=CD guide CD20 entry;
 CC WWW=http://www.ncbi.nlm.nih.gov/prov/cd/cd20.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X12530; CAA31046.1; -
 DR EMBL: M27394; AAA35581.1; -
 DR EMBL: M27395; -; NOT ANNOTATED_CDS.
 DR EMBL: L23419; AAA88911.1; JOINED.
 DR EMBL: L23415; AAA88911.1; JOINED.
 DR EMBL: L23416; AAA88911.1; JOINED.
 DR EMBL: L23417; AAA88911.1; JOINED.
 DR EMBL: X07203; CAA30179.1; -
 DR EMBL: X07204; CAA30180.1; -
 DR PIR: A27400; A27400.
 DR PIR: J10042; J10042.
 DR PIR: A30586; A30586.
 DR PIR: S00387; S00387.
 DR MIM: 112210; -
 KM B-cell; Transmembrane; Phosphorylation.
 FT DOMAIN 1 63 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 64 84 POTENTIAL.
 FT TRANSMEM 85 105 POTENTIAL.
 FT TRANSMEM 121 141 POTENTIAL.

FT TRANSMEM 189 209 POTENTIAL.
 FT DOMAIN 210 297 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 81 167 PROBABLE.
 FT DISULFID 167 183 PROBABLE.
 FT DISULFID 111 220 PROBABLE.
 FT CONFLICT 13 13 P -> L (IN REF. 4).
 FT CONFLICT 71 71 M -> I (IN REF. 3).
 SQ SEQUENCE 297 AA; 33077 MW; AC5420F8B626BD1 CRC64;
 Query Match 14.38; Score 110.5; DB 1; Length 297;
 Best Local Similarity 28.88; Pred. NO. 0.0029;
 Matches 30; Conservative 21; Mismatches 46; Indels 7; Gaps 3;
 QY 34 STGSPLOKLFARKKKIIGTIOILGIMFSGVIFLTLKPYRRF-PFIFLSGYPTMG 92
 DB 36 SLVGPSTGSEFMRESKTTGAVQINGLPHIALGGL---LMPAGIYAVICVTWYPIWGG 91
 QY 93 VLFINSGLFIAVRKRTETIGILITMTFSEIEFLSPSIL 136
 DB 92 IMYIISGSLAATKRNKCL--YKGMIMNSLSLFAISGMIL 133
 RESULT 5
 FCER_HUMAN STANDARD; PRT; 244 AA.
 ID FCER_HUMAN
 AC 001362;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR BETA-SUBUNIT (FCER1) (IGE FC RECEPTOR, BETA-SUBUNIT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92316966; PubMed=1535625;
 RA Kuester H., Zhang L., Brihi A.T., Macglashan D.W., Kinet J.-P.;
 RT "The gene and cDNA for the human high affinity immunoglobulin E receptor beta chain and expression of the complete human receptor.";
 RL J. Biol. Chem. 267:12782-12787(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92339505; PubMed=1386024;
 RA Maekawa K., Imagawa N., Tanaka Y., Harada S.;
 RT "Determination of the sequence coding for the beta subunit of the human high-affinity IGE receptor.";
 RL FEBS Lett. 302:161-165(1992).
 RN [3]
 RP VARIANT GLU-237.
 RX MEDLINE=96414302; PubMed=8817330;
 RA Hall M.R., Cookson W.O.;
 RT "A new variant of the beta subunit of the high-affinity receptor for immunoglobulin E (Fc epsilon RI beta E237G): associations with measures of atopy and bronchial hyper-responsiveness.";
 RL Hum. Mol. Genet. 5:959-962(1996).
 RN [4]
 RP VARIANT GLU-237.
 RX MEDLINE=96440420; PubMed=8842731;
 RA Shirakawa T., Mao X.-Q., Sasaki S., Enomoto T., Kawai M., Morimoto K., Hopkin J.;
 RT "Association between atopic asthma and a coding variant of Fc-epsilon RI-beta in a Japanese population.";
 RL Hum. Mol. Genet. 5:1129-1130(1996).
 RN [5]
 RP ERRATUM.
 RX MEDLINE=97123518; PubMed=8968765;
 RA Shirakawa T., Mao X.-Q., Sasaki S., Enomoto T., Kawai M., Morimoto K., Hopkin J.;
 RL Hum. Mol. Genet. 5:2068-2068(1996).

```

CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN EPSILON. HIGH
CC AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC
CC RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL
CC ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)
CC RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR
CC ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
CC -1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO
CC DISULFIDE LINKED GAMMA CHAINS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: FOUND ON THE SURFACE OF MAST CELLS AND
CC BASOPHILS.
CC -1- POLYMORPHISM: VARIANT GLU-237 HAS BEEN FOUND TO BE PRESENT IN ABOUT
CC 5.3% OF A 1004 INDIVIDUALS POPULATION SAMPLE IN AUSTRALIA. IT
CC SEEMS TO BE A RISK FACTOR FOR ATOPIC DERMATITIS AND ASTHMA.
CC
CC CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC DR EMBL: D10583; BAA01440.1; -
CC DR EMBL: M89796; AAA60269.1; -
CC DR PIR: S21154; S21154.
CC DR PIR: A42806; A42806.
CC DR MIM: 147138; -
CC
CC KW IGE-binding protein; Receptor; Transmembrane; Polymorphism.
CC
CC FT DOMAIN 1 59
CC FT TRANSSEM 60 79
CC FT TRANSSEM 80 97
CC FT TRANSSEM 98 117
CC FT TRANSSEM 118 130
CC FT TRANSSEM 131 150
CC FT TRANSSEM 151 180
CC FT TRANSSEM 181 200
CC FT TRANSSEM 201 244
CC FT TRANSSEM 237 237
CC FT VARIANT 237
CC
CC SQ SEQUENCE 244 AA; 26533 MW; CE523102D5F567AF CRC64;

```

```

OX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VA2 / PYRO4 / CXXC3:
RX MEDLINE=89296483; PubMed=2662141;
RA Brown T.A., Constable A., Ray J.A., Waring R.B., Scanzocchio C.,
RA Davies R.W.;
RT "Nucleotide sequence of the Aspergillus nidulans mitochondrial gene
RT for subunit 5 of NADH dehydrogenase.";
RL Nucleic Acids Res. 17:4371-4371(1989).
RN [2]
RP SEQUENCE OF 316-349 FROM N.A.
RC STRAIN=NRRL 322;
RL Luo X., Khan N.Q., Wientjes F.J.;
RL Submitted (NOV-1991) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: NADH + UBIOQUINONE = NAD(+) + UBIOQUINOL.
CC
CC CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC DR EMBL: X15011; CAA33116.1; -
CC DR EMBL: X62993; CAA44727.1; -
CC DR PIR: S04724; S04724.
CC DR InterPro: IPR001516; -
CC DR InterPro: IPR001750; -
CC DR Pfam: PF00361; oxidored_q1; 1.
CC DR Pfam: PF00662; oxidored_q1_N; 1.
CC KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
CC
CC SQ SEQUENCE 657 AA; 73361 MW; B5B679A151E7287 CRC64;

```

Query Match 10.6%; Score 82; DB 1; Length 657;
 Best Local Similarity 26.8%; Pred. NO. 2.5;
 Matches 38; Conservative 27; Mismatches 39; Indels 38; Gaps 8;

```

QY 22 EYESTEL-----SAPTEST-----OSPLQKFAKKMILGIIQILFGIMTFSPVIF 68
DB 268 EYSTVVLVLCWLGATITVSSILGLDFQIIRKLFAYSTWQGMVIAVAGLSSTN--LA 324
QY 69 LFTLLKPYRPPFFELSGYFPGSVLFINSGAFIIVK-----RK--TTEITGILITLM 120
DB 325 LFHLV-----NNAFYKALLFLGAGSVIHAADVADNDPKRYGGLAEPLPIITYVM 372
QY 121 TFSIIEFLISLPSILGCHSED 142
DB 373 LIASLSL-VAVPF-MTGFSYMD 392

```

RESULT 7
 ID CAN1_YEAST STANDARD; PRT; 590 AA.
 AC P04817;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE ARGININE PERMEASE.
 GN CAN1 OR YEL063C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=88165106; PubMed=3327612;
 RA Ahmad M., Bussey H.;
 RT "Yeast arginine permease: nucleotide sequence of the CAN1 gene.";
 RL Curr. Genet. 10:587-592(1986).
 RN [2]

SEQUENCE FROM N.A.
RX MEDLINE-86008235; Pubmed-3900064;
RA Hoffmann W.;
RT "Molecular characterization of the CAN1 locus in *Saccharomyces cerevisiae*. A transmembrane protein without N-terminal hydrophobic signal sequence.";
RL J. Biol. Chem. 260:11831-11837(1985).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R., Aviles E., Bero A., Brennan T., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G., Hunkeler-Smith S., Hyman R., Kayser A., Komp C., Lashkari D., Law H., Lin D., Mosedale D., Nakahara K., Nemath A., Norgren R., Oelner P., Oh C., Patel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V., Taylor P., Wei Y., Yellon M., Botstein D., Davis R.W.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: HIGH-AFFINITY PERMEASE FOR ARGININE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: X03784; CAA27416.1; -
DR EMBL: M1724; AAA34467.1; -
DR EMBL: U18795; AAB65024.1; -
DR PIR: A23922; GRBYR.
DR SGD: S0000789; CAN1.
DR InterPro: IPR002027; -
DR Pfam: PF00324; aa_permeases; 1.
DR PROSITE: PS00218; AMINO ACID PERMEASE_1; 1.
KM Transport; Amino-acid transport; Transmembrane.
FT DOMAIN 1 92
FT TRANSMEM 93 110
FT TRANSMEM 230 254
FT TRANSMEM 324 343
FT TRANSMEM 377 396
FT TRANSMEM 420 437
FT TRANSMEM 451 469
FT TRANSMEM 494 518
FT TRANSMEM 525 548
FT DOMAIN 549 590
FT CONFICT 534 534
SEQUENCE 590 AA; 65785 MW; 4E5A21C77145330D CRC64;

Query Match 10.6%; Score 81.5; DB 1; Length 590;
Best Local Similarity 22.9%; Pred. No. 2.5;
Matches 27; Conservative 26; Mismatches 42; Indels 23; Gaps 4;

OY 20 ASYESTELSATFESQSPLOKLFARKMKILGTIOLFGIMTFSEGVIFLTLKPY--P 77
DB 294 ALFFQGTGLVGTAGEANRKRKSVPAIK-----KVFRLITYIGSLFGLVLPYNDP 348
OY 78 RPFIFLISGVFWGSLVFLNSGAFLLAVKKRTETLG-----LLITLMTFSIIEFLPI 129
DB 349 KLT-----QSTYVSTSPFIILAEISGKVLPHINAVILITLIISANSNIVY 396

RESULT 8
TCR_BACST
ID TCR_BACST STANDARD; PRT; 458 AA.
AC P07561;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)

DE TETRACYCLINE RESISTANCE PROTEIN.
GN TET.
OS *Bacillus stearothermophilus*.
OG Plasmid pTH15.
OC Bacteria; Firmicutes; *Bacillus*/Clostridium group;
OC *Bacillus*/staphylococcus group; *Bacillus*.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-86031344; Pubmed-2996983;
RA Hoshino T., Ikeda T., Tomizuka N., Furukawa K.;
RT "Nucleotide sequence of the tetracycline resistance gene of pTH15, a thermophilic *Bacillus* plasmid: comparison with staphylococcal *TCR* controls.";
RL Gene 37:131-138(1985).
CC -1- FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTI-PORTER.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSLUCASE FAMILY).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: M1036; AAA22851.1; -
DR PIR: A23973; YRBSRP.
DR InterPro: IPR001411; -
DR PRINTS: PR01036; TCRTETB.
KW Antibiotic resistance; Transmembrane; Transport; Symport; Plasmid.
FT TRANSMEM 12 33
FT TRANSMEM 81 100
FT TRANSMEM 111 128
FT TRANSMEM 140 162
FT TRANSMEM 165 185
FT TRANSMEM 201 221
FT TRANSMEM 223 240
FT TRANSMEM 256 276
FT TRANSMEM 297 317
FT TRANSMEM 324 344
FT TRANSMEM 346 365
FT TRANSMEM 432 451
SEQUENCE 458 AA; 50119 MW; E1F33FAC57E7B06A CRC64;

Query Match 10.4%; Score 80; DB 1; Length 458;
Best Local Similarity 25.8%; Pred. No. 2.7;
Matches 32; Conservative 21; Mismatches 41; Indels 30; Gaps 6;

OY 32 TFSQSPLOKLFARKMKILGTIOLFGIMTFSEGVIFLTLKPYRPFIFLISGVFWG 91
DB 178 TIITVPLMKLLKKEVRIKHPDK-GIILMSYIVF-----MLFTSYST-- 223
OY 92 SVLFNSGAFLLAVK--RKTE-----TVGLITLMTFSIIEFLISL-PSFI 135
DB 224 SFLIVSVLSFLVKKHRIKVTDFVDPGLGKNIPFMIGVLGCGIIFTVAGFVSMVPM 283
OY 136 LGCH 139
DB 284 KDVA 287

RESULT 9
TCR_STRAG
ID TCR_STRAG STANDARD; PRT; 458 AA.
AC P13924;
DT 01-JAN-1990 (Rel. 13, Created)

DT 01-JAN-1990 (Rel. 13, last sequence update)
 DT 01-OCT-1994 (Rel. 30, last annotation update)
 DE TETRACYCLINE RESISTANCE PROTEIN.
 GN TET.
 OS Streptococcus agalactiae.
 OC plasmid pMV158.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1311;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90016790; PubMed=2677995.
 RA van der Lelie D., Bron S., Venema G., Oskam L.;
 RT 'Similarity of minus origins of replication and flanking open reading
 frames of plasmids pUB110, pPB913 and pMV158.';
 RL Nucleic Acids Res. 17:7283-7294(1989).
 CC -1- FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE
 EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE
 ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN
 FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTI-PORTER.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
 AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X15669; CAA3712.1; -
 DR PIR: C25599; YTSOG.
 DR InterPro: IPR001411; -
 DR PRINTS: PR01036; TCRTEB.
 KW Antibiotic resistance; Transmembrane; Transport; Symport; Plasmid.
 FT TRANSMEM 12 33 POTENTIAL.
 FT TRANSMEM 81 100 POTENTIAL.
 FT TRANSMEM 111 129 POTENTIAL.
 FT TRANSMEM 140 162 POTENTIAL.
 FT TRANSMEM 165 185 POTENTIAL.
 FT TRANSMEM 201 221 POTENTIAL.
 FT TRANSMEM 223 240 POTENTIAL.
 FT TRANSMEM 256 276 POTENTIAL.
 FT TRANSMEM 297 317 POTENTIAL.
 FT TRANSMEM 324 344 POTENTIAL.
 FT TRANSMEM 346 365 POTENTIAL.
 FT TRANSMEM 432 451 POTENTIAL.
 SQ SEQUENCE 458 AA; 50006 MW; AD2014E7CA19995 CRC64;

Query Match 10.4%; Score 80; DB 1; Length 458;
 Best Local Similarity 25.8%; Pred. No. 2.7;
 Matches 32; Conservative 21; Mismatches 41; Indels 30; Gaps 6;

OY 32 TFSQSLQRLFAARKMLICTIQLFGIMTSEFVFLFTLLKRPYPRPFILSGYPFWG 91
 DB 178 TITVPLMKLTKKEVRIKGFEDIK-GILMSYGVFE-----MLFTTSYSL-- 223
 OY 92 SVLEFNGALLIAVK--RKTE-----TLGILITLMTSEITLFLISL-PEST 135
 DB 224 SFLIVSVLSFLIFVKHIRKYTDEVPDGLKNILFMIGVLCGGIIFVAGFVSMVPM 283
 OY 136 LGCH 139
 DB 284 KDVA 287

RESULT 10
 TCR_STRPN
 ID TCR_STRPN STANDARD; PRT; 458 AA.
 AC P11063; P72219;

DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, last sequence update)
 DT 15-DEC-1998 (Rel. 37, last annotation update)
 DE TETRACYCLINE RESISTANCE PROTEIN.
 GN TET.
 OS Streptococcus pneumoniae; Bacillus cereus; and Bacillus subtilis.
 OG plasmid pLS1, plasmid pBC16, plasmid pHY163PLK, plasmid pPB19, and
 OG plasmid pNS1981.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313, 1396, 1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX SPECIES=S.pneumoniae; PLASMID=pLS1;
 RC MEDLINE=87226167; PubMed=2438417;
 RA Lacks S.A., Lopez P., Greenberg B., Espinosa M.;
 RT 'Identification and analysis of genes for tetracycline resistance and
 replication functions in the broad-host-range plasmid pLS1.';
 RL J. Mol. Biol. 192:753-765(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=B.cereus; PLASMID=pBC16;
 RX MEDLINE=90221899; PubMed=2109312;
 RA Paiva A., Vidgren G., Simonen M., Rintala H., Laamanen P.;
 RT 'Nucleotide sequence of the tetracycline resistance gene of pBC16
 from Bacillus cereus.';
 RL Nucleic Acids Res. 18:1635-1635(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC PLASMID=pHY163PLK;
 RA Ishiwa H., Shibahara H.;
 RT 'New shuttle vectors for Escherichia coli and Bacillus subtilis. III.
 Nucleotide sequence analysis of tetracycline resistance gene of
 pnp1. J. Genet. 60:485-498(1985).
 RL [4]
 RP SEQUENCE FROM N.A.
 RC PLASMID=pHY163PLK;
 RX MEDLINE=83129391; PubMed=6186390;
 RA Selzer G., Som T., Itoh T., Tomizawa J.;
 RT 'The origin of replication of plasmid pLSA and comparative studies on
 the nucleotide sequences around the origin of related plasmids.';
 RL Cell 32:119-129(1983).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC SPECIES=B.subtilis; PLASMID=pPB19;
 RX MEDLINE=92052681; PubMed=1946749;
 RA Oskam L., Hillenga D.J., Venema G., Bron S.;
 RT 'The large Bacillus plasmid pPB19 contains two integrated
 rolling-circle plasmids carrying mobilization functions.';
 RL Plasmid 26:30-39(1991).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC SPECIES=B.subtilis; STRAIN=168 / MABBURG; PLASMID=pNS1981;
 RX MEDLINE=86287699; PubMed=3090576;
 RA Sakaguchi R., Shishido K., Hoshino T., Furukawa K.;
 RT 'The nucleotide sequence of the tetracycline resistance gene of
 plasmid pNS1981 from Bacillus subtilis differs from pPRT15 from a
 thermophilic Bacillus by two base pairs.';
 RL Plasmid 16:72-73(1986).
 CC -1- FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE
 EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE
 ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN
 FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTI-PORTER.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
 AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial

```

RT      XIV from Saccharomyces cerevisiae between the BNI1 and the POL2
RT      genes."
RL      Yeast 12.5:505-514(1996).
CC      -1- FUNCTION: HIGH-AFFINITY PERMEASE FOR BASIC AMINO ACIDS.
CC      -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC      -1- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X74069; CAA52199.1; -
DR      EMBL; X92494; CAA63228.1; -
DR      EMBL; Z71546; CAA96177.1; -.
DR      PIR; S44329; S44329.
DR      SGD; S0005214; ALP1.
DR      InterPro; IPR002027; -
DR      Pfam; PF00324; aa-permeases; 1.
DR      PROSITE; PS00218; AMINO ACID PERMEASE 1; 1.
KW      Transport; Amino-acid transport; Transmembrane.
FT      TRANSMEM 75 95 POTENTIAL.
FT      TRANSMEM 96 116 POTENTIAL.
FT      TRANSMEM 154 174 POTENTIAL.

```

FT	TRANSMEM	266	286	POTENTIAL.
FT	TRANSMEM	307	327	POTENTIAL.
FT	TRANSMEM	360	380	POTENTIAL.
FT	TRANSMEM	404	424	POTENTIAL.
FT	TRANSMEM	432	452	POTENTIAL.
FT	TRANSMEM	478	498	POTENTIAL.
FT	TRANSMEM	511	531	POTENTIAL.
FT	CONFLICT	51	51	D -> D (IN REF. 1).
FT	CONFLICT	126	126	V -> A (IN REF. 1).
FT	CONFLICT	260	260	D -> N (IN REF. 1).
FT	CONFLICT	517	517	I -> V (IN REF. 1).
FT	CONFLICT	548	548	R -> H (IN REF. 1).
SO	SEQUENCE	573 AA;	64013 MW;	359DPEI166C348A1 CRC64;

Query Match	10.28;	Score 78.5;	DB 1;	Length 573;
Best Local Similarity	24.18;	Pred. NO. 4.5;		
Matches 28;	Conservative 24;	Mismatches 45;	Indels 19;	Gaps 4;

QY	20	ASEYESTELSATPSTSTOSPLQKLFARKKKTIGTIOILEGIMTFESGVILEFTLLKPYPRF	79
Db	277	AFYQGTETLVGVTAGEANAPRKALPRAIK-----KVVRILVFIYILSDFGLVPPY-ND	330
QY	80	PFILSGVPFPGSGSLFINSAGFLTAVRKRTETELG-----ILTIMFSTIELEFI	129
Db	331	PKLSDG-----TFVSSSPFMSTIENGSKVLPIDFINAVVLLITLSAGNSNYI	379

RESULT 12				
ANSP_SALTY				
ID	ANSP_SALTY	STANDARD;	PRT;	497 AA.
AC	P40812;			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, last sequence update)			
DT	01-FEB-1995 (Rel. 31, last annotation update)			
DE	L-ASPARAGINE PERMEASE (L-ASPARAGINE TRANSPORT PROTEIN).			
GN	ANSP.			
OS	Salmonella typhimurium.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Salmonella.			
OX	NCBI_taxid=602;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
CC	STRAIN-SA2656;			

RX MEDLINE=95202072; PubMed=7894705;
 RA Jennings P.J., Anderson J.K., Beacham I.R.;
 RT "Cloning and molecular analysis of the *Salmonella enterica* ansp gene,
 RT encoding an L-asparagine permease.";
 RL Microbiology 141:141-146(1995).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U04851; AAA80001.1; -;
 DR StyGene; SG10527; ansp.
 DR InterPro: IPR002027; -;
 DR Pfam: PF00324; aa-permeases; 1.
 DR PROSITE: PS00218; AMINO ACID PERMEASE 1; 1.
 KM Transport; Amino-acid transport; Transmembrane; Inner membrane.
 FT TRANSMEM 34 54 POTENTIAL.
 FT TRANSMEM 58 78 POTENTIAL.
 FT TRANSMEM 109 129 POTENTIAL.
 FT TRANSMEM 146 166 POTENTIAL.
 FT TRANSMEM 171 191 POTENTIAL.
 FT TRANSMEM 219 239 POTENTIAL.
 FT TRANSMEM 264 284 POTENTIAL.
 FT TRANSMEM 298 318 POTENTIAL.
 FT TRANSMEM 353 373 POTENTIAL.
 FT TRANSMEM 378 398 POTENTIAL.
 FT TRANSMEM 422 442 POTENTIAL.
 FT TRANSMEM 448 468 POTENTIAL.
 SO SEQUENCE 497 AA; 54004 MW; BE80CB117992CBBB CRC64;
 Query Match 10.18; Score 77.5; DB 1; Length 497;
 Best Local Similarity 22.8%; Pred. No. 4.9;
 Matches 37; Conservative 25; Mismatches 47; Indels 53; Gaps 8;
 Y 23 YESTELSATFTSTQSPLOKLFARKMKILGTQLIFGIMTFSFGVIFLTLKPRPRF--- 79
 Db 233 FASTELVGTAGAECKDPQKMKPKAIN---SVIMRIGLFYGVSVLVLLPLPNNAYQAG 287
 Y 80 --PFI-FLS--GYPFMGSVLF-----NSG-----AFLLAVK 106
 Db 288 QSPVTFEFSKRGVYIGSIMNIVVLTAAALSLNGLYCTGRIILKSMGSGAPFMAKMS 347
 Y 107 RKTETUGILITLMTFSEILFLIS-----LPPSLIG 137
 Db 348 ROHVYAGIILATLVV-VGVFELNYLPSRVEIYLVNFSIG 388
 RESULT 13
 COTE_HUMAN STANDARD; PRT; 669 AA.
 AC P81408;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE COTE1 PROTEIN.
 GN COTE1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hippocampus;
 RA MEDLINE=97474796; PubMed=9331372;
 RA Winfield S.L., Tayebi N., Martin B.M., Glins E.I., Sidransky E.;
 RT "Identification of three additional genes contiguous to the

RT glucocerebrosidase locus on chromosome 1q21: implications for Gaucher
 RT disease.";
 RL Genome Res. 7:1020-1026(1997).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.
 CC -1- SIMILARITY: TO HUMAN KIA0574.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF023268; AAC51822.1; -;
 DR Transmembrane.
 KM TRANSMEM 34 54 POTENTIAL.
 FT TRANSMEM 67 87 POTENTIAL.
 FT TRANSMEM 91 111 POTENTIAL.
 FT TRANSMEM 174 194 POTENTIAL.
 FT TRANSMEM 244 250 POLY-PRO.
 FT DOMAIN 635 639 POLY-SER.
 SO SEQUENCE 669 AA; 71482 MW; E30360AC9A4571B6 CRC64;
 Query Match 10.08; Score 77; DB 1; Length 669;
 Best Local Similarity 22.6%; Pred. No. 7.2;
 Matches 31; Conservative 25; Mismatches 67; Indels 14; Gaps 5;
 Y 16 PETTASEYESTELSATFTSTQSPLOKLFARKMKILGTQLIFGIMTFSFGVIFLTLK 74
 Db 3 PPSDDSSRLTSRSPSTGTLRLRHPMLQALLTLGLVQLVGLVYTFSMVSSVTTE 62
 Y 75 PYPRFPFIPLSGYFPGSVLFINSGLF-LVAVKRTETETGILITLMTFSEILFLISLP 133
 Db 63 SIKR-----SCPSWAGFSLAFSGVGVIVSWKRPFLVIFPSFL--SVLCVMISWAG 112
 Y 134 SILGCHSEDC--DCEQC 148
 Db 113 SVLSCKNAQLARDPQC 129
 RESULT 14
 COTE_HUMAN STANDARD; PRT; 669 AA.
 AC P29388;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 5 (EC 1.6.5.3).
 GN ND5 OR NAD5.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Mitochondrion.
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, COLUMBIA;
 RX MEDLINE=92007799; PubMed=1915303;
 RA Knoop V., Schuster W., Wissinger B., Brennicke A.;
 RT "Trans splicing integrates an exon of 22 nucleotides into the nad5
 RT mRNA in higher plant mitochondria.";
 RL EMBO J. 10:3483-3493(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, COLUMBIA;
 RX MEDLINE=97141919; PubMed=8988169;
 RA Unseid M., Marlenfeld J.R., Brandt P., Brennicke A.;
 RT "The mitochondrial genome of Arabidopsis thaliana contains 57 genes
 RT in 366,924 nucleotides.";

RL Nat. Genet. 15:57-61(1997).
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X60045; CAA42648.1; -
CC EMBL: X60047; CAA42648.1; JOINED.
CC EMBL: X60048; CAA42648.1; JOINED.
CC EMBL: Y08501; CAA69752.1; -
CC PIR: S20234; DNM005.
CC Mendel: 2011; ARATH; nad5;1.
CC InterPro: IPR001516; -
CC InterPro: IPR001750; -
CC Pfam: PF00361; oxidored_q1; 1.
CC Pfam: PF00662; oxidored_q1; 1.
CC Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
CC SEQUENCE 669 AA; 73907 MW; 0748DCB81DB40429 CRC64;
CC -----
Query Match 9.8%; Score 75.5; DB 1; Length 669;
Best Local Similarity 24.3%; Pred. No. 9.9;
Matches 36; Conservative 30; Mismatches 51; Indels 31; Gaps 8;
QY 8 SPVFLVFPPE---ITASEYESTELSATFTSQSPLOKFLPA-RKMKILGTLIQLFGIMTF 62
DB 273 SPLF-EVPPALVITTSAGATTSFLATGILNDLKRVIAVSCSGLGIMTFACGINSY 331
QY 63 SPGYIFLTLKPPRPFFFLSGYPPWGSVLPINSGAFILAVK-----RK---TTETLG 114
DB 332 SVSFFHLM-----NHAFFKALFLSAGSVIHMSDEQMRKMGGLASSFP 376
QY 115 ILITMTFSIIEFLISLPSILGCHSD 142
DB 377 LTVAMMLIGSLSL-TGPPF-LTGFTSKD 402
CC -----
RESULT 15
ATCU_BACSU
ID ATCU_BACSU STANDARD; PRT; 803 AA.
AC 032220;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
POTENTIAL COPPER-TRANSPORTING ATPASE (EC 3.6.1.36).
YVYGX.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBL_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Kunst F.; Ogasawara N.; Yoshikawa H.; Danchin A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INVOLVED IN COPPER TRANSPORT (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + ORTHOPHOSPHATE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
CC (EI-E2 ATPASES). SUBFAMILY IB.
CC -1- SIMILARITY: CONTAINS 1 HEAVY-METAL-ASSOCIATED (HMA) DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z99121; CAB15355.1; -
CC DR HSSP: P04129; 2HQ1.
CC DR Subtilisin; BG14106; YVYGX.
CC DR InterPro: IPR001757; -
CC DR InterPro: IPR001934; -
CC DR Pfam: PF00122; EI-E2 ATPase; 1.
CC DR Pfam: PF00403; HMA; 2.
CC DR PROSITE: PS00154; ATPASE_EI_E2; 1.
CC DR PROSITE: PS01047; HMA; 2.
CC KW Hydrolyase; Transmembrane; Phosphorylation; ATP-binding; Copper;
CC KM Metal-binding; Repeat.
CC FT DOMAIN 12 41 HMA.
CC FT METAL 17 109 HMA.
CC FT METAL 17 17 COPPER (POTENTIAL).
CC FT METAL 20 20 COPPER (POTENTIAL).
CC FT METAL 85 85 COPPER (POTENTIAL).
CC FT METAL 88 88 COPPER (POTENTIAL).
CC FT TRANSMEM 163 183 POTENTIAL.
CC FT TRANSMEM 197 217 POTENTIAL.
CC FT TRANSMEM 229 249 POTENTIAL.
CC FT TRANSMEM 260 280 POTENTIAL.
CC FT TRANSMEM 416 436 POTENTIAL.
CC FT TRANSMEM 448 468 POTENTIAL.
CC FT TRANSMEM 610 630 POTENTIAL.
CC FT TRANSMEM 704 724 POTENTIAL.
CC FT TRANSMEM 767 787 POTENTIAL.
CC FT MOD_RES 500 500 POTENTIAL.
CC SQ SEQUENCE 803 AA; 86024 MW; D9C8DA5D40326C6B CRC64;
CC -----
Query Match 9.8%; Score 75.5; DB 1; Length 803;
Best Local Similarity 23.8%; Pred. No. 12;
Matches 40; Conservative 25; Mismatches 54; Indels 49; Gaps 10;
QY 6 AHSVFLVFPPEITASEYESTELSATFTSQSPLOKFLPA-RKMKILG----- 51
DB 103 ANAVY--NFALETYTVENRKEASVS--DLKEAVDKL-GYKLKLGKGDSEAAAKKKEER 157
QY 52 -TTQLIFGIMTFSGYI-----FLFT-----LLKPYPR-----PFIISGYPPWG 91
DB 158 KQTRALFSAV-LSFPLLMNAVSHFTFSFIWVPDIFLNPMPQALATPVQGLGWPPY- 215
QY 92 SVLFINSGAFILAVKRTTETGLITLMTFSIIEFLISLPSILGCH 139
DB 216 -----VGAY--KALRNKSNAMDVLAJGTAAAYASLYLTFOSIGSH 255

Search completed: October 7, 2001, 04:46:10
Job time: 365 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2001, 04:07:35 ; Search time 40.14 Seconds
(without alignments)
300.552 Million cell updates/sec

Title: US-09-735-712-2

Perfect score: 1015
Sequence: 1 MDSTAHSPVFLVFPPEITA.....ISLPSTLIGCHSDCCQEC 199

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

A.Geneseq_0601:*

- 1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:*
- 22: /SIDSL/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	294	29.0	77	21	AAAG03074
2	206.5	20.3	239	20	AAW96745
3	206.5	20.3	239	21	AAV50174
4	206.5	20.3	245	21	AAV94973
5	184.5	18.2	214	17	AAW06503
6	184.5	18.2	214	19	AAW41056
7	184.5	18.2	225	21	AAV94449
8	184.5	18.2	242	21	AAV73495
9	184.5	18.2	248	20	AAV15225
10	184.5	18.2	248	21	AAV91531
11	184.5	18.2	248	22	AAW70489

12	184.5	18.2	250	20	AAV48505	Human breast tumor
13	184.5	18.2	257	21	AAW58419	Lung cancer associ
14	184.5	18.2	273	21	AAV91680	Human secreted pro
15	176.5	17.4	267	20	AAV15224	Human receptor pro
16	174.5	17.2	299	21	AAV91352	Human secreted pro
17	170	16.7	297	13	AAW20808	Human CD20 antigen
18	169	16.7	248	20	AAV36046	Extended human sec
19	168	16.6	297	10	AAV91356	CD20.4 antigen
20	168	16.6	297	21	AAV96131	Human cell surface
21	167	16.5	297	17	AAW91436	Human CD20 antigen
22	167	16.5	297	19	AAW80445	Human CD20.4 anti
23	167	16.5	297	20	AAW81352	Human CD20.4 anti
24	162.5	16.0	237	22	AAW4585	Human HAIRERbs iso
25	162.5	16.0	238	22	AAW70490	Human HAIRERbs iso
26	162.5	16.0	240	21	AAV69998	Human HAIRERbs pro
27	162	16.0	239	22	AAW4584	Human receptor-ass
28	162	16.0	249	22	AAW74583	Human HAIRERbs iso
29	161.5	15.9	307	21	AAV91421	Human secreted pro
30	159	15.7	243	11	AAW05026	Beta subunit of ra
31	159	15.7	243	14	AAW42341	Beta subunit of the
32	159	15.7	246	12	AAW44770	Beta subunit of hl
33	159	15.7	246	14	AAW42337	Human FCER1 beta
34	147.5	14.5	250	19	AAW61619	Clone HTPPE86 of T
35	147.5	14.5	250	21	AAW70432	Human cell surface
36	147.5	14.5	302	21	AAW54312	Human pancreatic c
37	142.5	14.0	167	21	AAV53632	A bone marrow secr
38	133	13.1	204	21	AAV91396	Human secreted pro
39	124.5	12.3	151	19	AAW54438	Mouse novel secret
40	124.5	12.3	151	21	AAW10245	Murine adult splee
41	123.5	12.2	244	18	AAW29149	Human high affinit
42	123.5	12.2	244	19	AAW75918	Human beta subunit
43	123.5	12.2	244	22	AAW74447	Human wild-type FC
44	123.5	12.2	244	22	AAW72900	Human IGERB SEQ ID
45	121.5	12.0	192	19	AAW74885	Human secreted pro

ALIGNMENTS

RESULT 1	
AAAG03074	standard; Protein; 77 AA.
ID	AAAG03074:
AC	AAAG03074:
DT	06-OCT-2000 (first entry)
XX	Human secreted protein, SEQ ID NO: 7155.
DE	Human secreted protein, SEQ ID NO: 7155.
XX	Human: 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW	gene therapy; chromosome mapping.
XX	
OS	Homo sapiens.
XX	
PN	EP1033401-A2.
XX	
PD	06-SEP-2000.
XX	
PF	21-FEB-2000; 2000EP-0200610.
XX	
PR	26-FEB-1999; 99US-0122487.
XX	
PA	(GEST) GENSET.
XX	
PI	Dumas Malne Edwards J, Duclert A, Giordano J;
XX	
DR	WPI: 2000-500381/45.
XX	
DR	N-FSDB; AAC03080.
XX	
PT	New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT	obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT	diagnostic, forensic, gene therapy and chromosome mapping procedures -

PF	15-DEC-1998;	98DS-0213389.
XX		
PR	21-AUG-1997;	97US-0916902.
XX		
PA	(INCY-)	INCYTE PHARM INC.
PI	Bandman O, Corley NC, Lal P;	
DR	WPI: 2000-012123/01.	
XX	N-PSDB; AA232842.	
XX		
PS	Claim 1, Fig 1, 29pp: English.	
CC		
CC	This sequence represents human high affinity immunoglobulin E (IgE)	
CC	receptor-like protein (IGERB). Nucleic acids encoding IGERB were	
CC	initially identified in a brain cDNA library, this sequence being a	
CC	consensus. An allergic response is initiated by release of IgE	
CC	from B lymphocytes. The IgE molecules then bind to the high affinity	
CC	IgE receptor (FcεR1) present on mast cells and basophils, which triggers	
CC	the release of histamine and proteases from cytoplasmic granules and	
CC	leads to the synthesis of effectors of the allergic and inflammatory	
CC	response, such as prostaglandins, leukotrienes and cytokines. As IGERB	
CC	binds IgE, it may be administered to stimulate allergic and immune	
CC	responses in patients in whom IGERB is under expressed or inactive and	
CC	to supplement the patients own production of the protein. IGERB may also	
CC	be used as an antigen for the production of antibodies and to identify	
CC	candidate agonists and antagonists of IGERB expression and activity.	
CC	The antibodies may also be used in diagnosis. Antibodies and antagonists	
CC	may be administered to downregulate IGERB activity and reduce the	
CC	potency of inflammatory and allergic responses. They may be used in this	
CC	way to treat inflammatory disorders such as multiple sclerosis,	
CC	osteoarthritis, asthma and some complications of cancer. Conversely, the	
CC	agonists may be used to enhance immune responses.	
SO	Sequence	239 AA;
Query Match	20.3%;	Score 206.5; DB 21; Length 239;
Best Local Similarity	31.8%;	Pred. No. 2,1e-16;
Matches	49; Conservative	32; Mismatches 70; Indels 3; Gaps 2
OY	39	LOKLFAR-KMKLTGTOIFGIMTFSGVYIEFTLLKPYRPPFFTLGSGYPPMGSYLFTN 97
DB	54	Iqeklkgepkvlygvvqlltalsismqitumcmasnitygnpsivlylgtlswymfil 113 ,
OY	98	SCAFPIAVKRRKTELLILSRIMNLSALRAIAGIILMF--GELIDQWYICGSHQNSQ 155
	114	sgslsaaagirtcklvrgslgmaltstsvlaasgallinfslsalfsfhnpymcnygnsm 173
OY	156	KRAVTVFLGLDITLMTFSIIEFLTSLPSSIGC 189
DB	174	chgtnslmgldgmwlllsvlefciavlsatfgc 207
RESULT	4	
ID	AAAY94973	
AC	AAAY94973 standard; Protein: 245 AA.	
XX		
DT	16-JUN-2000	(first entry)
DE	Human secreted protein clone pe246_1 protein sequence SEQ ID NO:152.	
XX		
KW	Human, secreted protein; immunestimulant; immunosuppressant; virucide;	
KW	antibacterial; antifungal; cytosstatic; antiinflammatory; dermatological;	
KW	antidiabetic; antisthmatic; antiarthritic; antirheumatic; protozoacide;	
KW	antihyroid; immune deficiency; severe combined immunodeficiency; SCID;	
KW	infection; HIV; hepatitis; malaria; autoimmune disorder; systemic lupus;	
KW	connective tissue disease; multiple sclerosis; erythematosis;	

KW Rheumatoid arthritis; autoimmune pulmonary inflammation; asthma;
 KM Guillain-Barre syndrome; autoimmune thyroiditis; myasthenia gravis;
 KM insulin dependent diabetes mellitus; graft-versus-host-disease;
 KM autoimmune inflammatory eye disease; allergy.
 XX Homo sapiens.
 OS
 XX
 PN MO200009552-A1.
 PD
 XX 24-FEB-2000.
 PF 13-AUG-1999; 99WO-US18298.
 XX
 PR 14-AUG-1998; 98US-0096622.
 PR 17-AUG-1998; 98US-0096815.
 PR 04-SEP-1998; 98US-0099229.
 PR 23-OCT-1998; 98US-0105368.
 PR 08-JAN-1999; 99US-0115234.
 PR 12-FEB-1999; 99US-0119931.
 PR 18-FEB-1999; 99US-0120575.
 PR 30-APR-1999; 99US-0132020.
 PR 11-AUG-1999; 99US-0096622.
 XX
 PA (GENEX) GENETICS INST INC.
 PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Agostino MJ, Steinhinger RJ, Spaulding V;
 PI Wong GS, Clark HF, Fechtel K;
 DR WPI: 2000-205979/18.
 XX
 XX New polynucleotides encoding secreted proteins, which may have e.g.
 PT nutritional, chemokine, immune stimulating or suppressing,
 PT hematopoiesis regulating, tissue growth, activin/inhbln
 PT antiinflammatory or tumor inhibition activity
 PS
 PS Claim 161; Page 607; 641pp: English.
 XX
 XX AAA16618 to AAA16697 encode the human secreted proteins given in
 CC CCAAY94898 to AAY94980, isolated from human adult brain, adult thyroid,
 CC adult retina, foetal carcinoma, adult blood, adult neural, foetal kidney,
 CC adult placenta, adult testis, whole embryo, adult cartilage, kidney,
 CC foetal brain, adult thymus, foetal placenta, adult uterus, adult tumour,
 CC and adult bladder, cDNA libraries. The polynucleotides and proteins are
 CC predicted to have biological activities which would make them suitable
 CC for treating, preventing or ameliorating medical conditions in humans
 CC and animals. The polynucleotides can be used as markers for tissues in
 CC which the protein is preferentially expressed, as molecular weight
 CC markers on Southern gels, and as chromosome markers or tags to identify
 CC chromosomes or to map gene positions. The proteins can be used in the
 CC treatment of immune deficiencies and disorders, such as severe combined
 CC immunodeficiency (SCID), as well as viral, bacterial, fungal and other
 CC infections. These infections include human immunodeficiency virus (HIV),
 CC hepatitis, herpesviruses, mycobacteria, Leishmania spp., malaria and
 CC candidiasis. The proteins can be used to treat autoimmune disorders such
 CC as connective tissue disease, multiple sclerosis, systemic lupus
 CC erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation,
 CC Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent
 CC diabetes mellitus, myasthenia gravis, graft-versus-host-disease and
 CC autoimmune inflammatory eye disease. The proteins can also be used to
 CC treat allergic conditions, such as asthma. AAA16698 to AAA16774 represent
 CC probes for the human secreted proteins from the present invention.
 CC
 XX
 XX Sequence 245 AA:
 1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 3

QY	98	SCAFIIAIVKRTETETIIISRIINLLSRAIAGIILLTF--GEIILDQNYICGRSHQNSQ	155
Db	120	sgslsaagirttkglvrsIgmicsvYlaasgillntislafsfthpvcnygnsm	179
QY	156	CKAAVTVLFGLIITLTMTFSIIEFLISLPSILGC	189
Db	180	chgtmsllmgdmvllslvlefcIavslsaafgc	213
RESULT	5		
ID	AAW06503	standard; Protein; 214 AA.	
XX	AAW06503;		
XX	06-FEB-1997	(first entry)	
DT	HTm4 protein.		
XX	Human HTm4 protein; Fc(epsilon)RI receptor; FR; atopic disease; allergy; asthma; atopic dermatitis; allergic rhinitis; hereditary.		
XX	Homo sapiens.		
XX	Key	Location/Qualifiers	
FT	Domain	/note= "Transmembrane domain"	
FT	Domain	/note= "Transmembrane domain"	
FT	Domain	/note= "Transmembrane domain"	
FT	Domain	/note= "Transmembrane domain"	
FT	Domain	/note= "Transmembrane domain"	
XX	US5552312-A.		
PN	03-SEP-1996.		
PD	06-OCT-1994;	94US-0318492.	
XX	06-OCT-1994;	94US-0318492.	
XX	(BETH-) BETH ISRAEL HOSPITAL BOSTON.		
PA	Adra-CN, Lelias J, Lim B;		
PL	WPI; 1996-412066/41.		
DR	N-PSDB; AAT45120.		
XX	New isolated mammalian HT(M4) protein DNA - used to develop prods. for use in the diagnosis and treatment of atopic diseases, e.g. allergy or asthma		
PT	Claim 2; Fig 1; 10pp; English.		
PS	This sequence represents human HTm4 protein. HTm4 is related to the Fc(epsilon)RI receptor (FR). HTm4 contains four hydrophobic domains of 20-21 amino acids. The N-terminal region before the first hydrophobic domain contains four proline residues, and each of the hydrophilic regions between the hydrophobic regions contains one proline. This protein crosses the plasma membrane four times forming two small extracellular loops and has both the N- or C-terminals in the cytoplasm. The HTm4 gene and protein can be used in the research and study of the induction of expression of FR and the function of its subunits. They can be used in the design of drugs which can block or inhibit induction of FR, thereby treating atopic diseases such as allergy, asthma, atopic dermatitis and allergic rhinitis. The HTm4 gene sequences may also be used to detect patients suffering from hereditary atopic disorders and to map genes on the human chromosome.		
XX	Sequence	214 AA;	

Query Match	Score 184.5;	DB 17;	Length 214;
Best Local Similarity	28.2%;	Pred. No. 7.3e-14;	
Matches	59;	Conservative 39;	Mismatches 86; Indels 25; Gaps 8;
Qy	1	MDSSTAH-SPVELVPEPETTASEYESTELSATFTSTOSPLQKLFARKMKILGTLQILFCI	59
Db	11	lgsasahbtgpgsetgpeelntsvyh-----plngsydyk--aklvgalgaqlnaa	60
Qy	60	MFFSGVIFLFLFKPY-----PRPFIFLSCYPFMGSVULFINSQAFILAVRKRTETILIT	115
Db	61	mla1gav-flgslgypynhfkhnffitytgyplwgaavfcssglsvaglkprtwlq	119
Qy	116	LSRIKMLLSALFAINGIILTFPGFILDONVI--CCYSHONSQ---CKAVYVFLGILIT	169
Db	120	nsfgmiasatalaygtatflsinlsvnqlsrsc---hsssespalcnymsgisngwsl	176
Qy	170	IMTFSIIEFLFISLPFSILGCHSEDCDCQ	198
Db	177	llllllelcvltistlamwncancnsre	205
RESULT 6			
AAW41056			
ID	AAW41056 standard; Protein: 214 AA.		
XX			
AC	AAW41056;		
XX			
DT	29-APR-1998 (first entry)		
XX			
DE	HTM4 protein.		
XX			
KM	HTM4; antibody; Ige Fc receptor; Fcepsilon1onribeta; CD20 antigen; TRAF-1;		
KM	4-transmembrane spanning protein superfamily; ligand binding mimic;		
KM	haematopoietic cell detection; inhibitor; tumour necrosis factor; KAP;		
KM	TNF receptor-associated factor; CDK-activating kinase; TRAF-2; TRAF-3;		
XX	TMASF.		
OS	Homo sapiens.		
XX			
PN	US5705615-A.		
XX			
PD	06-JAN-1998.		
XX			
PF	03-SEP-1996; 96US-0707340.		
XX			
PR	03-SEP-1996; 96US-0707340.		
XX			
PR	06-OCT-1994; 94US-0318492.		
XX			
PR	03-JUL-1996; 96US-0675648.		
XX			
PA	(BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.		
XX			
PI	Adra CN, Lelias J, Lam B;		
XX			
DR	WPI: 1998-086194/08.		
XX			
PI	N-PSDB; AAV03875.		
XX			
PT	Antibody specific for HTM4 protein - useful to block or mimic ligand		
XX	binding or detect haematopoietic cells		
PS	Disclosure: Column 25-28; 18pp; English.		
XX			
CC	This sequence is the HTM4 protein. The HTM4 protein is specifically		
CC	bound by the antibody of the invention. HTM4 shows homology to the beta		
CC	subunit of the high affinity Ige Fc receptor (Fcepsilon1onribeta) and the		
CC	CD20 antigen, both of which are in the 4-transmembrane spanning proteins		
CC	superfamily (TMASF). The antibody of the invention is used to block or		
CC	mimic binding of ligands such as the tumour necrosis factor (TNF)		
CC	receptor-associated factors TRAF-1, TRAF-2 and TRAF-3 and the		
CC	phosphatase, CDK-activating kinase (KAP), to receptors comprising HTM4,		
CC	and to detect haematopoietic cells.		

PT Novel polynucleotides and proteins having biological activities which
 PT make them suitable for treating, preventing or ameliorating medical
 PT conditions in humans or animals -
 XX
 XX Claim 221: Page 718; 730pp; English.
 XX
 CC The present invention describes human secreted proteins encoded by
 CC polynucleotides obtained from adult testes, foetal brain, adult brain,
 CC brain (foetal and adult), foetal kidney, adult spleen, and adult thymus
 CC cDNA libraries. The polynucleotides and proteins are predicted to have
 CC biological activities which would make them suitable for treating,
 CC preventing or ameliorating medical conditions in humans and animals.
 CC Suggested activities include nutritional activity, cytokine and cell
 CC proliferation/differentiation activity, immune stimulating (e.g. as
 CC vaccines) or suppressing activity, haematopoiesis regulating activity,
 CC tissue growth activity, activin/inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic and thrombolytic activity, receptor/
 CC ligand activity, anti-inflammatory activity, cadherin/tumour invasion
 CC suppressor activity, and tumour inhibition activity. The polynucleotides
 CC are also stated to be useful for gene therapy. Therapeutic compositions
 CC are also presently valuable for veterinary applications. AA252475 to
 CC AA252581 encode human secreted proteins, and AA73390 to AA73500
 CC represent human secreted proteins, given in the present invention.
 CC
 XX Sequence 242 AA:
 SQ
 Query Match 18.2%; Score 184.5; DB 21; Length 242;
 Best Local Similarity 26.4%; Pred. No. 8.9e-14;
 Matches 55; Conservative 40; Mismatches 90; Indels 23; Gaps 6;
 QY 1 MDSSTASPVFLVPPPEITASEYESTLSATFSTQSPLOKLFARKKKILGTTQILFGIM 60
 Db 1 mtsqpyneitvipsnvl--nfsqaekpeptngqdsikkhlaekivglitqlcgmm 58
 QY 61 TFSGCVFLFTLKP-YRPFPIFL-SCYPPWGSVLINSAGFLIAVKRKTETELILSR 118
 Db 59 vlsigillaasfsptitqvtstllnsayptlpgfflissglsistatekrtikilvns1 118
 QY 119 IMNLLSLRAIVAGIIL-----TFGFLDPN-----YICGSHON---SCKAV 159.
 Db 119 vgsllsalsalvgfllsvkqatlnpaslqceldknhiprsvsyfhdsltydcyta 178
 QY 160 TVLFLGILITLMTFSIIEFLSPFSIL 187
 Db 179 kaslagclsimlctlliefclavltavl 206
 RESULT 9
 AAY15225
 ID AAY15225 standard; protein: 248 AA.
 XX
 AC AAY15225:
 XX
 DT 26-OCT-1999 (first entry)
 XX
 DE Human receptor protein (HURP) 4 amino acid sequence.
 XX
 KM receptor; cancer; autoimmune disorder; inflammation;
 KM antagonist; cell surface protein; cell signalling;
 KM antibody; human receptor protein; HURP; reproductive disorder;
 KM developmental disorder; gastrointestinal disorder.
 XX
 OS Homo sapiens.
 XX
 FH key Location/Qualifiers
 FT Misc-difference 8 /note= "Potential N-glycosylation site"
 FT Misc-difference 20 /note= "Potential N-glycosylation site"
 FT Misc-difference 74 /note= "Potential N-glycosylation site"
 FT Misc-difference 110 /note= "Potential N-glycosylation site"

FT /note= "Potential CAMP-/cGMP-dependent protein-
 FT kinase phosphorylation site"
 FT Misc-difference 22 /note= "Potential casein kinase II-
 FT phosphorylation site"
 FT Misc-difference 193 /note= "Potential casein kinase II-
 FT phosphorylation site"
 FT Misc-difference 36 /note= "Potential protein kinase C-
 FT phosphorylation site"
 FT Misc-difference 105 /note= "Potential protein kinase C-
 FT phosphorylation site"
 FT Misc-difference 136 /note= "Potential protein kinase C-
 FT phosphorylation site"
 FT Misc-difference 177 /note= "Potential protein kinase C-
 FT phosphorylation site"
 FT Misc-difference 233 /note= "Potential protein kinase C-
 FT phosphorylation site"
 FT Misc-differences 242 /note= "Potential tyrosine kinase
 FT phosphorylation"
 XX
 PN W09941375-A2.
 XX
 PD 19-AUG-1999.
 XX
 PF 05-FEB-1999; 99WO-US02572.
 XX
 PR 12-FEB-1998; 98US-0022939.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Au-Young J, Bandman O, Baughn M, Corley NC, Guegler KJ;
 PI Hillman JL, Lai P, Shah P, Tang YT, Yae H;
 XX WPI: 1999-494536/41.
 DR N-PSDB; AA206369.
 XX
 PT New human receptor proteins, used e.g. to treat, prevent and
 PT diagnose gastrointestinal and developmental disorders - and related
 PT nucleic acids, vectors, transformed cells, antibodies, agonists and
 PT antagonists
 PS Claim 1; Page 81-82; 94pp; English.
 XX
 CC The human receptor protein 4 (HURP-4) has 22% homology with the
 CC rat IGE receptor and 19% homology with human CD20 protein.
 CC HURP-4 is expressed in cancerous, inflamed, hematopoietic/immune
 CC and gastrointestinal tissue. HURP-4 therefore appears to have a role in
 CC some forms of cancer, autoimmune/inflammatory disorders, and
 CC gastrointestinal disorders.
 CC This gives rise to the possibility of using an antagonist or an antibody
 CC of HURP to treat or prevent cancer or autoimmune/inflammatory disorders.
 XX
 SQ Sequence 248 AA:
 Query Match 18.2%; Score 184.5; DB 20; Length 248;
 Best Local Similarity 26.4%; Pred. No. 8.9e-14;
 Matches 55; Conservative 40; Mismatches 90; Indels 23; Gaps 6;
 QY 1 MDSSTASPVFLVPPPEITASEYESTLSATFSTQSPLOKLFARKKKILGTTQILFGIM 60
 Db 1 mtsqpyneitvipsnvl--nfsqaekpeptngqdsikkhlaekivglitqlcgmm 58
 QY 61 TFSGCVFLFTLKP-YRPFPIFL-SCYPPWGSVLINSAGFLIAVKRKTETELILSR 118
 Db 59 vlsigillaasfsptitqvtstllnsayptlpgfflissglsistatekrtikilvns1 118

OY 119 IKNLSALEAIGIILL-----TFGFIIDON-----YICGSHON---SOCKAV 159
 Db 119 vsiilsalsaiygfllsvkqatlmpasigceldkmiprtsrsvyryhdslyttcya 178
 OY 160 TVLFGILITMTFTSIELEFISHPESIL 187
 Db 179 kaslgatlsmltcltllefcclavtlavl 206
 RESULT 10
 ID AAY91531
 XX AAY91531 standard; Protein; 248 AA.
 AC AAY91531;
 XX
 DT 29-JUN-2000 (first entry)
 XX
 XX Human secreted protein sequence encoded by gene 81 SEQ ID NO:204.
 KW Human; secreted protein; diagnosis: cytostatic; immunosuppressive;
 KW antiHIV; antiinflammatory; nootropic; neuroprotective; antiallergic;
 KW osteopathic; antiarthritic; antibacterial; antidiabetic; antisthma;
 KW antipsoriatic; cardiant; gene therapy; cancer; neurological disorder;
 KW immune disease; inflammation; blood disorder; tumour; chromosome 1.
 XX
 OS Homo sapiens.
 XX
 PN WO200006698-A1.
 PD 10-FEB-2000.
 XX
 PF 29-JUL-1999; 99WO-US17130.
 XX
 PR 30-JUL-1998; 98US-0094657.
 PR 05-AUG-1998; 98US-0095486.
 PR 06-AUG-1998; 98US-0095454.
 PR 06-AUG-1998; 98US-0095455.
 PR 12-AUG-1998; 98US-0096319.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;
 PI Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;
 PI Soppet DR, Endress GA, Edner R, Olsen HS, Mucenski M;
 DR WPI: 2000-195282/17.
 DR N-PSDB: AAA26426.
 XX
 PS New isolated human genes and the secreted polypeptides they encode,
 PS useful for diagnosis and treatment of e.g. cancers, neurological
 PS disorders, immune diseases, inflammation or blood disorders
 PS
 PS Claim 11: Page 513-514; 634pp; English.
 CC The polynucleotide sequences given in AAA26346 to AAA26458 encode the
 CC human secreted proteins given in AAY91451 to AAY91691. The human secreted
 CC proteins can have activities based on the tissues and cells they are
 CC expressed in. Examples of the activities are: cytostatic;
 CC immunosuppressive; antiHIV; antiinflammatory; nootropic; neuroprotective;
 CC antiallergic; osteopathic; antiarthritic; antibacterial; antidiabetic;
 CC antisthma; antipsoriatic; and cardiant. The polynucleotides and their
 CC corresponding secreted proteins are useful for preventing, treating or
 CC ameliorating medical conditions, e.g. by protein or gene therapy. Also
 CC pathological conditions can be diagnosed by determining the amount of the
 CC proteins in a sample or by determining the presence of mutations in the
 CC polynucleotides. Specific uses are described for each of the
 CC polynucleotides, based on which tissues they are most highly expressed
 CC in, and include developing products for the diagnosis or treatment of
 CC cancer, tumours, neurodegenerative disorders, developmental abnormalities
 CC and foetal deficiencies, blood disorders, diseases of the immune system,
 CC autoimmune diseases, hepatic and renal disease, inflammation,
 CC allergies, Alzheimer's and behavioural disorders, schizophrenia,

	CC	osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
	CC	transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
	CC	cardiovascular disorders, reproductive disorders, gastrointestinal
	CC	disorders, respiratory disorders and metabolic disorders. The proteins
	CC	or polynucleotides can also be used as food additives or preservatives.
	CC	The proteins are also useful for identifying their binding partners.
	CC	AA26337 to AA26345 and AY91450 are sequences used in the
	CC	exemplification of the present invention.
	CC	
SQ	Sequence	248 AA;
	Query Match	18.2%; Score 184.5; DB: 21; Length 248;
	Best Local Similarity	26.4%; Pred. NO. 8.9e-14;
	Matches	55; Conservative 40; Mismatches 90; Indels 23; Gaps 6;
OY	1 MDSTASHSPVFLVPFPEPTATASEESTELSATFTFSQSPLKLFARKMILGTIQLTFCIM 60	: : : : : : : : : : : : : : : : :
Dd	1 mtsqvpnettiivlsnyl--nfsgaeekpepingqdskkhkbelkvigtliglgmm 58	: : : : : : : : : : : : : : : : :
OY	61 TFSGVIFLFLTKLP-YRPPETFL-SCYPEMGSVLTNSGAFLLAVKRKTETLIISR 118	: : : : : : : : : : : : : : : : :
Dd	59 vslgliliasasfspnfcvstllinsayptfgtffiflisgslstatetekrltklivhesl 118	: : : : : : : : : : : : : : : : :
OY	119 IWNLSALRAVAGITLL-----TFGRITDQN-----YICGYSHON---SCKAY 159	: : : : : : : : : : : : : : : : :
Dd	119 vgsllisaalsavglfllsvkgatlnpaslqceldkhnlprrsyvsyfyhdsltytcyta 178	: : : : : : : : : : : : : : : : :
OY	160 TVLFGILITLMTFSIIIEFLSLPSIL 187	: : : : : : : : : : : : : : : : :
Dd	179 kaslegxismiclllefciaavlavl 206	: : : : : : : : : : : : : : : : :
	RESULT 11	
AAB70489	ID AAB70489 standard; Protein: 248 AA.	
XX AC	AAB70489;	
XX XX	04-MAY-2001 (first entry)	
XX DE	Human hHAIERBs-Iso protein sequence SEQ ID NO:7.	
XX KW	Human; hHAIERBs-Iso; HAIERBs; HAIERBs isomer; detection;	
XX KW	high affinity immunoglobulin epsilon receptor beta subunit.	
XX OS	Homo sapiens.	
XX PN	CNI269410-A.	
XX PD	11-OCT-2000.	
XX PF	17-MAR-2000; 2000CN-0114959.	
XX PR	17-MAR-2000; 2000CN-0114959.	
XX PA	(SREN-) SOUTHERN RES CENT NAT HUMAN GENE GROUP.	
XX PI	Xiao H, Liu F, Song H;	
XX DR	WPI: 2001-050545/07.	
XX DR	N-PADB; AAF63724.	
XX PT	New human immunoglobulin receptor subunit protein and its nucleic acid	
XX PS	-	
XX PS	Claim 4; Page 17; 22pp; Chinese.	
CC CC	The present invention describes a human high affinity immunoglobulin	
CC CC	epsilon receptor beta subunit isomer, designated hHAIERBs-Iso.	
CC CC	hHAIERBs-Iso is isolated from in human pheochromocytoma. The present	
CC CC	invention also describes methods for the preparation and detection of	
CC CC	hHAIERBs-Iso protein and nucleotide sequences. The present sequence	

FT Misc-difference 13 /note= "Potential N-glycosylation site"
 FT Misc-difference 228 /note= "Potential N-glycosylation site"
 FT Misc-difference 232 /note= "Potential N-glycosylation site"
 FT Misc-difference 260 /note= "Potential N-glycosylation site"
 FT Misc-difference 8 /note= "Potential N-glycosylation site"
 FT Misc-difference 115 /note= "Potential casein kinase II-phosphorylation site"
 FT Misc-difference 118 /note= "Potential casein kinase II-phosphorylation site"
 FT Misc-difference 198 /note= "Potential casein kinase II-phosphorylation site"
 FT Misc-difference 3 /note= "Potential protein kinase C-phosphorylation site"
 FT Misc-difference 115 /note= "Potential protein kinase C-phosphorylation site"
 FT Misc-difference 198 /note= "Potential protein kinase C-phosphorylation site"

MO9941375-A2.
 19-AUG-1999.
 05-FEB-1999; 99MO-US02572.
 12-FEB-1998; 98US-0022939.

(INCYTE PHARM INC.

Au-Young J, Bandman O, Baughn M, Corley NC, Guegler KJ, Hillman JL, Lal P, Shah P, Tang YT, Yue H;

WPI: 1999-494536/41.
 N-PSDB; AAZ06368.

New human receptor proteins, used e.g. to treat, prevent and diagnose gastrointestinal and developmental disorders - and related nucleic acids, vectors, transformed cells, antibodies, agonists and antagonists

Claim 1; Page 81; 94pp; English.

The Human receptor protein 3 (HURP-3) has 23% homology with the rat IGE receptor beta sub-unit and 19% with human CD20. HURP-3 is expressed in cancerous, inflamed, foetal/proliferating, and reproductive tissue. HURP-3 therefore appears to have a role in some forms of cancer, autoimmune/inflammatory disorders, reproductive disorders, and developmental disorders. This gives rise to the possibility of using an antagonist or an antibody of HURP to treat or prevent cancer or autoimmune/inflammatory disorders.

Sequence 267 AA;

Query Match 17.4%; Score 176.5; DB 20; Length 267;
 Best Local Similarity 30.1%; Pred. No. 8,7e-13;
 Matches 44; Conservative 29; Mismatches 54; Indels 19; Gaps 3;

QY 43 PARKMKILGIIQIFGIMFSGVIFL---FTLIKPYRPFIFLSGYPFGSVLFTNSG 99
 DB 84 fkeekalgyiglmvghigfivclisfstrevigfastavlgypfwglsifisg 143
 QY 100 AFLIAVRKRTETLLILSRMTNLISALRATAGIILFFGILD---QNYICGSHONSQ 155
 DB 144 slsvsaskelstrclvsglsmivssilatifyilllvdmclngvagdywvavlsqk--- 200

QY 156 CKAATVLEGLILITMTSIIIEFIS 181
 DB 201 -----qisatlmfslleffva 217

Search completed: October 7, 2001, 04:39:26
 Job time: 1911 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2001, 04:07:40 ; Search time 24.66 Seconds
(without alignments)
166.159 Million cell updates/sec

Title: US-09-735-712-2

Perfect score: 1015

Sequence: 1 MDSTAHSPVFLVFPPEITA.....ISLPSTLGHSEDCCEQC 199

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCPTUS_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	206.5	20.3	239	2	US-08-916-902A-1 Sequence 1, App1
2	206.5	20.3	239	2	US-09-213-389-1 Sequence 1, App1
3	184.5	18.2	214	1	US-08-318-492-4 Sequence 4, App1
4	184.5	18.2	214	2	US-08-707-340-4 Sequence 4, App1
5	184.5	18.2	214	2	US-08-916-902A-3 Sequence 3, App1
6	184.5	18.2	214	2	US-08-994-578-4 Sequence 4, App1
7	184.5	18.2	214	2	US-09-213-389-3 Sequence 3, App1
8	159	15.7	243	1	US-07-869-933-29 Sequence 29, App1
9	159	15.7	243	1	US-08-201-879A-3 Sequence 33, App1
10	159	15.7	243	2	US-08-916-902A-4 Sequence 4, App1
11	159	15.7	243	2	US-08-201-879A-4 Sequence 4, App1
12	159	15.7	243	2	US-09-213-389-4 Sequence 29, App1
13	159	15.7	243	4	US-09-103-663-29 Sequence 33, App1
14	159	15.7	243	4	US-09-103-663-33 Sequence 23, App1
15	159	15.7	246	4	US-07-869-933-23 Sequence 33, App1
16	159	15.7	246	4	US-09-103-663-23 Sequence 33, App1
17	158.5	15.6	235	1	US-07-869-933-34 Sequence 34, App1
18	158.5	15.6	235	1	US-08-201-879A-5 Sequence 5, App1
19	158.5	15.6	235	4	US-09-103-663-34 Sequence 34, App1
20	123.5	12.2	244	1	US-07-869-933-32 Sequence 32, App1
21	123.5	12.2	244	4	US-08-201-879A-3 Sequence 32, App1
22	123.5	12.2	244	4	US-09-103-663-32 Sequence 32, App1
23	85	7.9	327	4	US-08-748-506-24 Sequence 24, App1
24	80.5	7.9	311	3	US-08-605-284B-23 Sequence 23, App1
25	80.5	7.9	1956	4	US-08-843-417-2 Sequence 1, App1
26	79.5	7.8	241	3	US-08-808-148-1 Patent No. 5386025
27	79	7.8	1872	6	5386025-6

28	79	7.8	1873	1	US-08-435-675B-4	Sequence 4, App1
29	78.5	7.7	423	2	US-08-494-907-14	Sequence 14, App1
30	78.5	7.7	423	5	PCT-US96-10986-14	Sequence 14, App1
31	78.5	7.7	956	2	US-08-897-443-3	Sequence 3, App1
32	78	7.7	347	4	US-09-097-889-14	Sequence 10, App1
33	77.5	7.6	1956	4	US-08-843-417-10	Sequence 4, App1
34	77	7.6	1416	1	US-08-061-465-4	Sequence 4, App1
35	76	7.5	1873	1	US-08-336-257A-7	Sequence 7, App1
36	75	7.4	459	4	US-08-097-889-22	Sequence 22, App1
37	73.5	7.2	381	1	US-08-467-125-2	Sequence 2, App1
38	73.5	7.2	381	2	US-08-911-320A-2	Sequence 2, App1
39	73.5	7.2	381	4	US-09-217-101-2	Sequence 2, App1
40	73.5	7.2	398	1	US-08-097-938-4	Sequence 4, App1
41	73.5	7.2	398	1	US-08-476-000-4	Sequence 4, App1
42	73.5	7.2	398	1	US-08-472-840-4	Sequence 4, App1
43	73.5	7.2	398	2	US-08-476-976-4	Sequence 4, App1
44	73.5	7.2	398	3	US-08-474-410-4	Sequence 4, App1
45	73	7.2	371	2	US-08-928-692-20	Sequence 20, App1

ALIGNMENTS

RESULT 1
US-08-916-902A-1
; Sequence 1, Application US/08916902A
; Patent No. 5871930
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
; TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; City: Palo Alto
; STATE: CA
; COUNTRY: USA
; Zip: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/916,902A
; FILING DATE: Herewith
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0371 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 239 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT04
; CLONE: 927955
; US-08-916-902A-1

Query Match	Similarity	20.3%	Score 206.5	DB 2	Length 239
Best Local	Similarity 31.8%	Pred No. 1.3e-15			
Matches 49	Conservative 32	Mismatches 70	Indels 3	Gaps 2	
OY	39	10KLEPAP-KMKILGTLQILFGIMTFSEGVYLFETLLKPYRPFPIFLSGYPWGSVYLFIN 97			
DB	54	10KEKFLKEGPKVLGVVQYDITLALMSISMGITMCMASNTGYSNPISVIGYITLINGSVMIIF 113			
OY	98	SCAPFLIAPKRTKTEVLIILSRIMNLSALRALAGIILTF--GFILIDWAYITGYSHNSQ 155			
DB	114	SSSLSIAAGIRTKTGKGLVRSISGMNTSSVLAASGLIINTFSIAFYSEFHHPCYNYGNSNN 173			
OY	156	CKAVTVLFLGILITLMTFSEIIEFLSLPISIGC 189			
DB	174	CHGTMSILMGIDGMVLLSLVLEFCIAVSLAAGC 207			
RESULT 2					
	09-213-389-1				
	Sequence 1, Application US/09213389				
	Patent No. 5977072				
	GENERAL INFORMATION:				
	APPLICANT: Bandman, Olga				
	APPLICANT: Lal, Preeti				
	APPLICANT: Corley, Neil C.				
	TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E				
	TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN				
	NUMBER OF SEQUENCES: 4				
	CORRESPONDENCE ADDRESS:				
	ADDRESSEE: Incyte Pharmaceuticals, Inc.				
	STREET: 3174 Porter Drive				
	CITY: Palo Alto				
	STATE: CA				
	COUNTRY: USA				
	ZIP: 94304				
	COMPUTER READABLE FORM:				
	MEDIUM TYPE: Diskette				
	COMPUTER: IBM Compatible				
	OPERATING SYSTEM: DOS				
	SOFTWARE: FASTSEQ for Windows Version 2.0				
	CURRENT APPLICATION DATA:				
	APPLICATION NUMBER: US/09/213.389				
	FILING DATE:				
	CLASSIFICATION:				
	PRIOR APPLICATION DATA:				
	APPLICATION NUMBER: 08/916.902				
	FILING DATE:				
	ATTORNEY/AGENT INFORMATION:				
	NAME: Billings, Lucy J.				
	REGISTRATION NUMBER: 36,749				
	REFERENCE/DOCKET NUMBER: PF-0371 US				
	TELECOMMUNICATION INFORMATION:				
	TELEPHONE: 415-855-0555				
	TELEFAX: 415-845-4166				
	TELEX:				
	INFORMATION FOR SEQ ID NO: 1:				
	SEQUENCE CHARACTERISTICS:				
	LENGTH: 239 amino acids				
	TYPE: amino acid				
	STRANDEDNESS: single				
	TOPOLOGY: linear				
	IMMEDIATE SOURCE:				
	LIBRARY: BRAINOT04				
	CLONE: 927955				
	US-09-213-389-1				
Query Match	20.3%	Score 206.5	DB 2	Length 239	
Best Local	Similarity 31.8%	Pred. No. 1.3e-15			
Matches 49	Conservative 32	Mismatches 70	Indels 3	Gaps 2	
OY	39	10KLEPAP-KMKILGTLQILFGIMTFSEGVYLFETLLKPYRPFPIFLSGYPWGSVYLFIN 97			
DB	54	10KEKFLKEGPKVLGVVQYDITLALMSISMGITMCMASNTGYSNPISVIGYITLINGSVMIIF 113			
OY	98	SCAPFLIAPKRTKTEVLIILSRIMNLSALRALAGIILTF--GFILIDWAYITGYSHNSQ 155			
DB	114	SSSLSIAAGIRTKTGKGLVRSISGMNTSSVLAASGLIINTFSIAFYSEFHHPCYNYGNSNN 173			
OY	156	CKAVTVLFLGILITLMTFSEIIEFLSLPISIGC 189			
DB	174	CHGTMSILMGIDGMVLLSLVLEFCIAVSLAAGC 207			
RESULT 2					
	09-213-389-1				
	Sequence 1, Application US/09213389				
	Patent No. 5977072				
	GENERAL INFORMATION:				
	APPLICANT: Bandman, Olga				
	APPLICANT: Lal, Preeti				
	APPLICANT: Corley, Neil C.				
	TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E				
	TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN				
	NUMBER OF SEQUENCES: 4				
	CORRESPONDENCE ADDRESS:				
</					

```

Db 54 LOEKRLKEPKVLGVQVOLLTALMSLSMGTMMCMASNTGNSPDISVIGYITLIMGSMFLII 113
OY 98 SCAPLIAVKKRTETPELLILSRIMNLISALRAIAGITLLTF--GFLIDONTICGYSQNSQ 155
Db 114 SGLSLIAGIRTKTKVLKMSLGMNTSVLASSGILINIFSLAFYSFHHYCNYSNNSN 173
OY 156 CKAATVLEFLGITLMTFTSIIELFISLPSISLIGC 189.
Db 174 CHGTMSIIMGIDGMVLLSVLEFCIAVLSLARFC 207

RESULT 3
US-08-318-492-4
Sequence 4, Application US/08318492
Patent No. 5552312
GENERAL INFORMATION:
APPLICANT: Lim, Bing
APPLICANT: Adra, Chaker N.
APPLICANT: Lellias, Jean-Michel
TITLE OF INVENTION: RECOMBINANT HTM4 GENE, PROTEIN AND
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,492
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: BIH94-03
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 214 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-318-492-4

Query Match 18.2%; Score 184.5; DB 1; Length 214;
Best Local Similarity 28.2%; Pred. No. 3,3e-13;
Matches 59; Conservative 39; Mismatches 86; Indels 25; Gaps 8,
OY 1 MDSSTAH-SPVFLVPPPELTASEYSTELSATFTSTQSLQKLFARKKIKLITQILRFI 59
Db 11 LGSASAHGTPESEGTGPEBELNLSVYH-----PIGSPDYOK--AKLQVIGAIIDLNAA 60
OY 60 MTFSGGVLEFLFLKPY-----PRPFIFLSGYPFGSVLEFINSGAFLIAVKKRTETLLI 115
Db 61 MILALGV-FLGSLQYPYHFKHFFFTTYTGCPIMGAVFECCSGTLSVAGIKPRITWQ 119
OY 116 LSRIMNLISALRAIAGITLLTFGFLIDONTYI--CGYSHQNSQ---CKAATVLEFLGIT 169
Db 120 NSFQGNINASATIALGTAFSLINIAVNTQSLRSC---HSSSESPDLQVNWGMSISNGWSL 176
OY 170 LMTFTSIIELFISLPSISLIGCHSEDDCQ 198
Db 177 LLITLLECVITSTIAMVCNANCNSRE 205

```

RESULT 4

US-08-707-340-4

; Sequence 4, Application US/08707340

; Patent No. 5705615

; GENERAL INFORMATION:

; APPLICANT: Lim, Bing

; APPLICANT: Adra, Chaker N.

; APPLICANT: LeJais, Jean-Michel

; TITLE OF INVENTION: RECOMBINANT HTM4 GENE, PROTEIN AND

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

; STREET: Two Militia Drive

; CITY: Lexington

; STATE: Massachusetts

; COUNTRY: U.S.A.

; ZIP: 02173

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/707,340

; FILING DATE: 03-SEP-1996

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/318,492

; FILING DATE: 06-OCT-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/675,648

; FILING DATE: 03-JUL-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Brook, David E.

; REGISTRATION NUMBER: 22,592

; REFERENCE/DOCKET NUMBER: B1H94-03A2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 861-6240

; TELEFAX: (617) 861-9540

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 214 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-707-340-4

Query Match 18.2%; Score 184.5; DB 1; Length 214;

Best Local Similarity 28.2%; Pred. No. 3.3e-13;

Matches 59; Conservative 39; Mismatches 86; Indels 25; Gaps 8;

QY 1 MDSSTAH-SVPLVFPPEITASEYESTELSATPTSTQSPLOKLFARKMKLITGLIQLFI 59

DB 11 LGSASAHGTGSGTEGPELMTSVYH-----PINGSPTYOK--AKLYVGAIGIQLINAA 60

QY 60 MTSFGVIFLFTLLKPY----PREPFIILSGYPWGSVLEFINSQAFLLAVKRTTETLLI 115

DB 61 MLALGV-FIGSLQYRPHQKHEFFFTYGYPIWGAVFCCSGTSLSVAGIKPTRTWIQ 119

QY 116 LSRMNLALRAIAGIILLTFGFIIDONYI--CGYSHQSQ-----CKATVLEFLGLIT 169

DB 120 NSFQMNIASTIALVGAFLSLNIAVNIQSLRSC--HSSSESPDLCLNYGSIISNGAVSL 176

QY 170 LMTFSIIEFLISLPFSLGCHSEDCDEQ 198

DB 177 LLITLLELCVTISTIAMMCNANCNSRE 205

US-08-916-902A-3

; Sequence 3, Application US/08916902A

; Patent No. 5871930

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; APPLICANT: Lal, Preeti

; APPLICANT: Corley, Neil C.

; TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incycle Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/916,902A

; FILING DATE: Herewith

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0371 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; TELEX:

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 214 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: GenBank

; CLONE: 561639

US-08-916-902A-3

Query Match 18.2%; Score 184.5; DB 2; Length 214;

Best Local Similarity 28.2%; Pred. No. 3.3e-13;

Matches 59; Conservative 39; Mismatches 86; Indels 25; Gaps 8;

QY 1 MDSSTAH-SVPLVFPPEITASEYESTELSATPTSTQSPLOKLFARKMKLITGLIQLFI 59

DB 11 LGSASAHGTGSGTEGPELMTSVYH-----PINGSPTYOK--AKLYVGAIGIQLINAA 60

QY 60 MTSFGVIFLFTLLKPY----PREPFIILSGYPWGSVLEFINSQAFLLAVKRTTETLLI 115

DB 61 MLALGV-FIGSLQYRPHQKHEFFFTYGYPIWGAVFCCSGTSLSVAGIKPTRTWIQ 119

QY 116 LSRMNLALRAIAGIILLTFGFIIDONYI--CGYSHQSQ-----CKATVLEFLGLIT 169

DB 120 NSFQMNIASTIALVGAFLSLNIAVNIQSLRSC--HSSSESPDLCLNYGSIISNGAVSL 176

QY 170 LMTFSIIEFLISLPFSLGCHSEDCDEQ 198

DB 177 LLITLLELCVTISTIAMMCNANCNSRE 205

RESULT 6

US-08-994-578-4

; Sequence 4, Application US/08994578

Patent No. 5972688
GENERAL INFORMATION:
APPLICANT: Lim, Bing
APPLICANT: Adra, Chaker N.
APPLICANT: Leillas, Jean-Michel
TITLE OF INVENTION: HTM4 METHODS OF TREATMENT AND ASSAYS,
TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/994,578
FILING DATE: December 19, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/707,340
FILING DATE: 03-SEP-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/675,648
FILING DATE: 03-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/318,492
FILING DATE: 06-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: BIH94-03A22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 861-6240
TELEFAX: (781) 861-9540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 214 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-994-578-4

Query Match 18.2%; Score 184.5; DB 2; Length 214;
Best Local Similarity 28.2%; Pred. No. 3.3e-13;
Matches 59; Conservative 39; Mismatches 86; Indels 25; Gaps 8;

QY 1 MDSSTAH-SPEVLVFPPEITASESESTELSATFTSTQSPLOKLFARKMKILGTQILFGI 59
DB 11 LGSASAHGTGSGETGPELMTSVH-----PINGSPTYOK--AKLOYGAIQIILNAA 60
QY 60 MTESEGVIFFTLLKPY-----PREPFIILSGYPFGWSVLFINSGAFILAVKRRKTTETLLI 115
DB 61 MILALGV-FLGSLOYRPHQKHFFFTFYGYPIWGAVFCSSGTLISVAGIKPTRWIO 119
QY 116 LSRIMNLSALRAIAGIILLTFGFIIDONYI--CGYSHONSQ---CKAVTVLFLGILIT 169
DB 120 NSFGMINASATIALVGAFLSLINAVNIQSLRSC---HSSESPEDLCLNYGSIISNGVSL 176
QY 170 LMTFSIIEFLISLPSILGCHSEDCDEQ 198
DB 177 LLITLLELCVTISTIAMCMNANCNSRE 205

RESULT 7
US-09-213-389-3
Sequence 3, Application US/09213389

Patent No. 5977072
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/213,389
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/916,902
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0371 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 214 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 561639
US-09-213-389-3

Query Match 18.2%; Score 184.5; DB 2; Length 214;
Best Local Similarity 28.2%; Pred. No. 3.3e-13;
Matches 59; Conservative 39; Mismatches 86; Indels 25; Gaps 8;

QY 1 MDSSTAH-SPEVLVFPPEITASESESTELSATFTSTQSPLOKLFARKMKILGTQILFGI 59
DB 11 LGSASAHGTGSGETGPELMTSVH-----PINGSPTYOK--AKLOYGAIQIILNAA 60
QY 60 MTESEGVIFFTLLKPY-----PREPFIILSGYPFGWSVLFINSGAFILAVKRRKTTETLLI 115
DB 61 MILALGV-FLGSLOYRPHQKHFFFTFYGYPIWGAVFCSSGTLISVAGIKPTRWIO 119
QY 116 LSRIMNLSALRAIAGIILLTFGFIIDONYI--CGYSHONSQ---CKAVTVLFLGILIT 169
DB 120 NSFGMINASATIALVGAFLSLINAVNIQSLRSC---HSSESPEDLCLNYGSIISNGVSL 176
QY 170 LMTFSIIEFLISLPSILGCHSEDCDEQ 198
DB 177 LLITLLELCVTISTIAMCMNANCNSRE 205

RESULT 8
US-07-869-933-29
Sequence 29, Application US/07869933
Patent No. 5770396
GENERAL INFORMATION:


```

1      REGISTRATION NUMBER: 36,749
2      REFERENCE/DOCKET NUMBER: PF-0371 US
3      TELECOMMUNICATION INFORMATION:
4      TELEPHONE: 415-855-0555
5      TELEFAX: 415-845-4166
6      TELEX:
7      INFORMATION FOR SEQ. ID NO.: 4:
8      SEQUENCE CHARACTERISTICS:
9          LENGTH: 243 amino acids
10         TYPE: 'amino acid
11         STRANDEDNESS: single
12         TOPOLOGY: linear
13         IMMEDIATE SOURCE:
14         LIBRARY: GenBank
15         CLONE: 204117
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012

```

Query Match	15.7%	Score 159;	DB 2;	Length 243;
Best Local Similarity	28.0%;	Pred. No. 2.9e-10;		
Matches 45;	Conservative 29;	Mismatches 71;	Indels 16;	Gaps 3

[illegible]

RESULT 12
US-09-213-389-4
Sequence 4, Application US/09213389
Patent No. 5977072
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTED for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/213,389
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/916,902
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0371 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:

```

; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 243 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; IMMEDIATE SOURCE:
;   LIBRARY: GenBank
;   CLONE: 204117
US-09-213-389-4

```

```

Query Match          15.7%; Score 159; DB 2; Length 243;
Best Local Similarity 28.0%; Pred. No. 2.9e-10;
Matches 45; Conservative 29; Mismatches 71; Indels 16; Gaps 3;

```

```

QY 36 QSPLOKLFARKMKILGTIQLFGIMFSGVIFLFTLLKP--YPRPFIFLSGYPFGSV 93
DB 47 QQTWQSLKKELEFLGVYVGLICFCFTVVCSTLQTSDEPDDEVLLYRAGYPPFGAV 106
94 LFNSGAFLLAVRKRTETILLIRIMNLSALRAIAGIILLTFGGFLDQNYICGSHON 153
107 LFVLSGFLSIMSRKNTLYVRGSLGANIVSSIAAGLAIILLNLSNNSAYM----- 159
QY 154 SOCKAVT-----VFLGLITLTMFPISTELFSLPFSIL 187
DB 160 NYCKDITEDDGCFTVSTITELVLMLEFLTLAFCSAVLLII 200

```

```

RESULT 13
US-09-103-663-29
; Sequence 29; Application US/09103663D
; Patent No. 6171803
; GENERAL INFORMATION:
;   APPLICANT: Kinet et al.
;   TITLE OF INVENTION: Isolation, characterization, and use of the human beta
;   TITLE OF INVENTION: subunit of the high affinity receptor for
;   FILE REFERENCE: 50490
;   CURRENT APPLICATION NUMBER: US/09/103,663D
;   EARLIER APPLICATION NUMBER: 07/869,933
;   EARLIER FILING DATE: 1992-04-16
;   NUMBER OF SEQ ID NOS: 35
;   SOFTWARE: PatentIn Ver. 2.1
;   SEQ ID NO 29
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   ORGANISM: Homo sapiens
09-103-663-29

```

```

Query Match          15.7%; Score 159; DB 4; Length 243;
Best Local Similarity 28.0%; Pred. No. 2.9e-10;
Matches 45; Conservative 29; Mismatches 71; Indels 16; Gaps 3;

```

```

QY 36 QSPLOKLFARKMKILGTIQLFGIMFSGVIFLFTLLKP--YPRPFIFLSGYPFGSV 93
DB 47 QQTWQSLKKELEFLGVYVGLICFCFTVVCSTLQTSDEPDDEVLLYRAGYPPFGAV 106
94 LFNSGAFLLAVRKRTETILLIRIMNLSALRAIAGIILLTFGGFLDQNYICGSHON 153
107 LFVLSGFLSIMSRKNTLYVRGSLGANIVSSIAAGLAIILLNLSNNSAYM----- 159
QY 154 SOCKAVT-----VFLGLITLTMFPISTELFSLPFSIL 187
DB 160 NYCKDITEDDGCFTVSTITELVLMLEFLTLAFCSAVLLII 200

```

```

RESULT 14
US-09-103-663-33
; Sequence 33; Application US/09103663D
; Patent No. 6171803

```

```

; GENERAL INFORMATION:
;   APPLICANT: Kinet et al.
;   TITLE OF INVENTION: Isolation, characterization, and use of the human beta
;   TITLE OF INVENTION: subunit of the high affinity receptor for
;   FILE REFERENCE: 50490
;   CURRENT APPLICATION NUMBER: US/09/103,663D
;   EARLIER APPLICATION NUMBER: 07/869,933
;   EARLIER FILING DATE: 1992-04-16
;   NUMBER OF SEQ ID NOS: 35
;   SOFTWARE: PatentIn Ver. 2.1
;   SEQ ID NO 33
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Rattus sp.
US-09-103-663-33

```

```

Query Match          15.7%; Score 159; DB 4; Length 243;
Best Local Similarity 28.0%; Pred. No. 2.9e-10;
Matches 45; Conservative 29; Mismatches 71; Indels 16; Gaps 3;
QY 36 QSPLOKLFARKMKILGTIQLFGIMFSGVIFLFTLLKP--YPRPFIFLSGYPFGSV 93
DB 47 QQTWQSLKKELEFLGVYVGLICFCFTVVCSTLQTSDEPDDEVLLYRAGYPPFGAV 106
94 LFNSGAFLLAVRKRTETILLIRIMNLSALRAIAGIILLTFGGFLDQNYICGSHON 153
107 LFVLSGFLSIMSRKNTLYVRGSLGANIVSSIAAGLAIILLNLSNNSAYM----- 159
QY 154 SOCKAVT-----VFLGLITLTMFPISTELFSLPFSIL 187
DB 160 NYCKDITEDDGCFTVSTITELVLMLEFLTLAFCSAVLLII 200

```

```

RESULT 15
US-07-869-933-23
; Sequence 23; Application US/07869933
; Patent No. 5770396
; GENERAL INFORMATION:
;   APPLICANT: KINET, Jean-Pierre
;   TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
;   TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
;   NUMBER OF SEQUENCES: 34
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Foley & Lardner
;   STREET: 1800 Diagonal Road, Suite 500
;   CITY: Alexandria
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22313-0299
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: floppy disk
;   OPERATING SYSTEM: IBM PC compatible
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/869,933
;   FILING DATE: 19920416
;   CLASSIFICATION: 514
;   ATTORNEY/AGENT INFORMATION:
;   NAME: BENT, Stephen A.
;   REGISTRATION NUMBER: 29,768
;   REFERENCE/DOCKET NUMBER: 40399/154 NIND
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (703)836-9300
;   TELEFAX: (703)683-4109
;   TELEX: 899149
;   INFORMATION FOR SEQ ID NO: 23:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 246 amino acids

```

TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-869-933-23

Query Match 15.78; Score 159; DB 1; Length 246;
Best Local Similarity 28.0%; Pred. No. 2.9e-10;
Matches 45; Conservative 29; Mismatches 71; Indels 16; Gaps 3;

QY 36 OSPLOKLFARRMKITIGTIOILFGIMTFSPGVIFLETLKP--YRPPPIFLSGYPPMGSV 93
Db 50 QQTWQSFLEKLELELVGVTVGLICFGIVVCSLTQTSDEDEVLLLYRAGYPEWGAV 109
QY 94 LFINSAGFLIAVKRRKTEETLIISRLNLSALRAIAGIILITFGFIIDQNYICGYSHON 153
Db 110 LFVLSGFLSIMSEKRNLYLVRSLSGANIVYSIAAGIAIILILNLSNNSAYM----- 162
154 SOCKAVT-----VFLGLITIMTFSITELFISLPSPSII 187
163 NYCKDITEDDGCFTVSFITELVMLFLFTLIAFCSAVILLII 203

Search completed: October 7, 2001, 04:40:00
Job time: 1940 sec

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	466.2	77.3	508	16	AI149899	AI149899 gfa3h06.x
2	456.6	74.7	516	7	AA436088	AA436088 zuo3a08.r
3	347	57.5	387	7	AA416972	AA416972 zt94h05.s
4	340.2	56.4	382	7	AA435988	AA435988 zuo3a08.s
5	291.8	48.4	394	1	AA758635	AA758635 ah67b04.s
6	283.8	47.1	484	163	BE107659	BE107659 UT-R-BT1
7	267.4	44.3	410	7	AA411806	AA411806 zt67a03.s
8	245.4	40.7	389	11	AA781801	AA781801 at58g01.s
9	242.8	40.3	415	7	AA470059	AA470059 zt94h05.r
10	207	34.3	482	14	AI002083	AI002083 ct38b02.s
11	141	23.4	332	147	BF319786	BF319786 uy63h11.x
12	131.8	21.9	601	138	BE638317	BE638317 EST00003
13	126.2	20.9	224	138	BE638335	BE638335 EST00002
14	115.6	19.2	538	224	AQ108532	AQ108532 CIT-HSP-2
15	113	18.7	281	11	AA707529	AA707529 ah41a12.s
16	86.4	14.3	1071	106	AL544561	AL544561 AL544561
17	84	13.9	793	154	BG484817	BG484817 602505511
18	83	13.8	218	138	BE638321	BE638321 EST000013
19	82.2	13.6	704	155	BG571626	BG571626 602592934
20	80.4	13.3	508	4	AA234138	AA234138 zt51b06.r
21	80.4	13.3	562	7	AA418443	AA418443 zv92e05.r
22	70.6	11.7	421	257	B86842	B86842 RPC111-26F1
23	61.2	10.1	887	106	AL531049	AL531049 AL531049
24	59.4	9.9	715	155	BG538851	BG538851 602568296
25	55.6	9.2	579	136	BE513276	BE513276 601315340
26	55.6	9.2	947	132	AK008652	AK008652 Mus muscu
27	55.4	9.2	470	102	AI807884	AI807884 w43g11.x
28	55	9.1	528	188	BE675149	BE675149 7103d02.x
29	55	9.1	604	172	BG024663	BG024663 602275469
30	54.6	9.1	997	143	BF056859	BF056859 7K10g10.x
31	54.4	9.0	1504	132	AK003110	AK003110 Mus muscu
32	53	8.8	1184	132	AK017928	AK017928 Mus muscu
33	52.6	8.7	515	153	BG434337	BG434337 602506331
34	52.6	8.7	594	32	AV717594	AV717594 AV717594
35	52.6	8.7	597	32	AV716310	AV716310 AV716310
36	52.6	8.7	619	32	AV661743	AV661743 AV661743
37	52.6	8.7	680	155	BG546745	BG546745 602574115
38	52.6	8.7	735	32	AV715678	AV715678 AV715678
39	52.4	8.7	504	150	BF522968	BF522968 UT-R-C2P-
40	52.4	8.7	755	134	AI950360	AI950360 wp10b06.x
41	52.2	8.7	736	137	BE563951	BE563951 601348160
42	52.2	8.7	860	106	AL551751	AL551751 AL551751
43	51.4	8.5	534	151	BF591089	BF591089 7h52g02.x
44	51.4	8.5	559	19	AI391038	AI391038 mcl10h04.y
45	50.8	8.4	482	151	BF600732	BF600732 265501 MA

ALIGNMENTS

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
1	AI149899/c	508 bp mRNA	EST	10-NOV-1998						
	gfa3h06.x1 Soares-testis_NHT Homo sapiens cDNA clone IMAGE:1752827									
	3' similar to SW_CD20_HUMAN P11836 B-LYMPHOCTE ANTIGEN CD20 ;									
	mRNA sequence.									
	AI149899									
	AI149899.1 GI:3678368									
	EST.									
	human.									
	Homo sapiens									
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.									
	1 (bases 1 to 508)									
	NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.									
	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),									
	Tumor Gene Index									

JOURNAL COMMENT

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-rt@mail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/dbp/image/image.html
Insert Length: 742 Std Error: 0.00
Seq primer: -40m13 fwd. Ef from Amersham
High quality sequence stop: 455.
Location/Qualifiers

FEATURES

Source
1..508
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1752827"
/clone_1ib="Soares-testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGGCCGCCCAATTTTCTTTTCTTTT 3']
Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73D vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT

186 a 91 c 89 g 142 t

ORIGIN

Query Match
Best Local Similarity 77.3%; Score 466.2; DB 16; Length 508;
Matches 468; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	133	agaaaaatgaatcttagagcatccagatcctgttggaattatgaccttctt	192
DB	508	agaaaaatgaatcttagagcatccagatcctgttggaattatgaccttctt	449
QY	193	gaagtatctcttcccttccctgtttaaaccatccagagttcccttattctt	252
DB	448	gaagtatctcttcccttccctgtttaaaccatccagagttcccttattctt	389
QY	253	taagatatccatctcgggctctgtttgttcattatctcggagccttctaatgca	312
DB	388	taagatatccatctcgggctctgtttgttcattatctcggagccttctaatgca	329
QY	313	gtgaaagaagaaacacagaaactcgtataatttagccgaataatgaatcttact	372
DB	328	gtgaaagaagaaacacagaaactcgtataatttagccgaataatgaatcttact	269
QY	373	gccctagagcaatagctgaatcattccctccatcattgtgttcattccatgataaac	432
DB	268	gccctagagcaatagctgaatcattccctccatcattgtgttcattccatgataaac	209
QY	433	tacatttgtgttattcaccacaaatagtcagtgtaagctgttactgtctgtcttg	492
DB	208	tacatttgtgttattcaccacaaatagtcagtgtaagctgttactgtctgtcttg	149
QY	493	ggaatttgattacattgttgaattcattacattatgattatcttctcgccttc	552
DB	148	ggaatttgattacattgttgaattcattacattatgattatcttctcgccttc	89
QY	553	tcatttggtggtccaccacagagattgtgtaacaaatgtgttga 603	
DB	88	tcatttggtggtccaccacagagattgtgtaacaaatgtgttga 38	

RESULT

2

THIS PAGE BLANK (USPTO)

AA436088 516 bp mRNA EST 09-NOV-1997
 LOCUS znu03a08.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:730742
 DEFINITION 5' similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ;
 mRNA sequence.
 ACCESSION AA436088
 VERSION AA436088.1 GI:2141002
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 516)
 AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Gelsel, G., Jost, S.,
 Kitzman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin
 White, Y., Wylie, T., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
 Mashu-NCI human EST Project
 Unpublished (1997)
 CONTACT: Wilson RK
 JOURNAL Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewatson.wustl.edu
 This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 733 Std Error: 0.00
 Seq primer: -28ml3 rev2 ET from Amersham
 High quality sequence stop: 482.
 FEATURES
 SOURCE
 1. 516
 /organism="Homo sapiens"
 /db_xref="GDB:5927437"
 /db_xref="taxon:9606"
 /clone="IMAGE:730742"
 /clone_1db="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech Laboratories
 , Inc., and primed with a Not I - oligo(dT) primer [5'
 TGTACCAATCTGAGTGGAGGAGCGGCCGCCCAATTTTATTTTATTTT 3']
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT7T3 vector. Library
 went through one round of normalization to Cot5, and was
 constructed by Bento Soares and M. Fatima Bonaldo. "
 BASE COUNT 149 a 117 c 80 g 170 t
 ORIGIN
 Query Match 74.7%; Score 450.6; DB 7; Length 516;
 Best Local Similarity 99.1%; Pred. No. 4.9e-105;
 Matches 453; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 atgagttcaagcacgcacacagtcggtgttctgtgtattcttcagaatactgct 60
 Db 52 ATGAGTTCAAGCACGCACACAGTCGCGTGTCTGTGATTTCTCCAGAAATCAGT 111
 QY 61 tcagaatattgtgtccacagaaatttcagcacgacccctttcaactcaagccctgcga 120
 Db 112 TCAGAAATTTGTGAGTCCACGAACTTTCAGCCAGACTTTTCACTCAAAAGCCCTTCA 171
 QY 121 aaattattgtctagaaaataaatactttagagacttccagatccgtgttggaattatg 180
 Db 172 AAATTTATTTGCTAGAAAATATCAAAATCTTAGGAGCTATCCAGATCGTGTGGAAATATG 231
 QY 181 acccttctcttggagttatcttcttccactgttaaacatatcaaggttcc 240
 Db 232 ACCCTTCTCTTGGAGTATCTCTTTTCACTTGTGTTAAACATATCCAGATTTCC 291

QY 241 ttatatattcttcagagatccattctctgggctctgttttcttaattcttgagcc 300
 Db 292 TTATATTCTTTCAGAGATATCATCTCTGGGCTCTGTGTTGTTCAATTATCTCGAGCC 351
 QY 301 ttccatttcagtggaagaaacacacagaaactctgtaatttgagccgaatratg 360
 Db 352 TTCTTATTTTCAGTGAAGAAAGAAACACAGAACTCTGTATATATGACCCAAATATG 411
 QY 361 aactctttagtgccctgagagcaatagctggaatcattctccacatttggttcac 420
 Db 412 AATTCTTGTAGTCCCTCGGAGCAATAGCTGAATCATCTCTCAATTTGTTTCATC 471
 QY 421 ctgattcaaacacacatttgggtttattcttaccaca 457
 Db 472 CTGATCAAAACATCAATTTGTGTTATTCACAAAA 508
 RESULT 3
 AA416972/c
 LOCUS zt94h05.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:730041
 DEFINITION 3' similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ;
 mRNA sequence.
 ACCESSION AA416972
 VERSION AA416972.1 GI:2077080
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 387)
 AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Gelsel, G., Jost, S.,
 Kitzman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin
 White, Y., Wylie, T., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
 Mashu-NCI human EST Project
 Unpublished (1997)
 CONTACT: Wilson RK
 JOURNAL Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewatson.wustl.edu
 This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 656 Std Error: 0.00
 Seq primer: -41m13 fwd ET from Amersham.
 FEATURES
 SOURCE
 1. 387
 /organism="Homo sapiens"
 /db_xref="GDB:5926570"
 /db_xref="taxon:9606"
 /clone="IMAGE:730041"
 /clone_1db="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech Laboratories
 , Inc., and primed with a Not I - oligo(dT) primer [5'
 TGTACCAATCTGAGTGGAGGAGCGGCCGCCCAATTTTATTTTATTTT 3']
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT7T3 vector. Library
 went through one round of normalization to Cot5, and was
 constructed by Bento Soares and M. Fatima Bonaldo. "
 BASE COUNT 138 a 77 c 67 g 105 t
 ORIGIN
 Query Match 57.5%; Score 347; DB 7; Length 387;
 Best Local Similarity 98.6%; Pred. No. 1.5e-78;
 Matches 350; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

THIS PAGE BLANK (USPTO)